

# UNIFORM DISTRIBUTION OF SUBPOLYNOMIAL FUNCTIONS ALONG PRIMES AND APPLICATIONS

VITALY BERGELSON, GRIGORI KOLESNIK, AND YOUNGHWAN SON

**ABSTRACT.** Let  $H$  be a Hardy field (a field consisting of germs of real-valued functions at infinity that is closed under differentiation) and let  $f \in H$  be a subpolynomial function. Let  $\mathcal{P} = \{2, 3, 5, 7, \dots\}$  be the (naturally ordered) set of primes. We show that  $(f(n))_{n \in \mathbb{N}}$  is uniformly distributed mod 1 if and only if  $(f(p))_{p \in \mathcal{P}}$  is uniformly distributed mod 1. This result is then utilized to derive various ergodic and combinatorial statements which significantly generalize the results obtained in [BKMST].

## 1. INTRODUCTION

In the recent paper [BKMST] new results on sets of recurrence involving the prime numbers were established, thereby unifying and refining some previous results obtained in [Sa1], [Sa2], [Sa3], [F], [KM] and [BL].

Here are the formulations of some of the results from [BKMST] that are relevant to the discussion in this introduction.

**Theorem 1.1** (Theorem 3.1 in [BKMST]). *Let  $c_1, c_2, \dots, c_k$  be distinct positive real numbers such that  $c_i \notin \mathbb{N}(\mathbb{N} := \{1, 2, 3, \dots\})$  for  $i = 1, 2, \dots, k$ . Let  $U_1, \dots, U_k$  be commuting unitary operators on a Hilbert space  $\mathcal{H}$ . Then,*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N U_1^{[p_n^{c_1}]} \dots U_k^{[p_n^{c_k}]} f = f^*,$$

*strongly in  $\mathcal{H}$ , where  $p_n$  denotes the  $n$ -th prime and  $f^*$  is the projection of  $f$  on  $\mathcal{H}_{inv} := \{f \in \mathcal{H} : U_i f = f \text{ for all } i\}$ .*

**Theorem 1.2** (Corollary 3.1 in [BKMST]). *Let  $c_1, c_2, \dots, c_k$  be positive non-integers. Let  $T_1, T_2, \dots, T_k$  be commuting, invertible measure preserving transformations on a probability space  $(X, \mathcal{B}, \mu)$ . Then, for any  $A \in \mathcal{B}$  with  $\mu(A) > 0$ , one has*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mu(A \cap T_1^{-[p_n^{c_1}]} \dots T_k^{-[p_n^{c_k}]} A) \geq \mu^2(A),$$

*where  $p_n$  denotes the  $n$ -th prime.*

Recall that the upper Banach density of a set  $E \subset \mathbb{Z}^k$  is defined to be

$$d^*(E) = \sup_{\{\Pi_n\}_{n \in \mathbb{N}}} \limsup_{n \rightarrow \infty} \frac{|E \cap \Pi_n|}{|\Pi_n|},$$

where the supremum is taken over all sequences of parallelepipeds

$$\Pi_n = [a_n^{(1)}, b_n^{(1)}] \times \dots \times [a_n^{(k)}, b_n^{(k)}] \subset \mathbb{Z}^k, \quad n \in \mathbb{N} \text{ with } b_n^{(i)} - a_n^{(i)} \rightarrow \infty \quad (1 \leq i \leq k).$$

---

The first author gratefully acknowledges the support of the NSF under grant DMS-1162073.

**Theorem 1.3** (Corollary 3.2 in [BKMST]). *Let  $c_1, \dots, c_k$  be positive non-integers. If  $E \subset \mathbb{Z}^k$  with  $d^*(E) > 0$ , then there exists a prime  $p$  such that  $([p^{c_1}], \dots, [p^{c_k}]) \in E - E$ . Moreover,*

$$\liminf_{N \rightarrow \infty} \frac{|\{p \leq N : ([p^{c_1}], \dots, [p^{c_k}]) \in E - E\}|}{\pi(N)} \geq d^*(E)^2.$$

To obtain combinatorial corollaries, such as Theorem 1.3, from the ergodic statements one utilizes the  $\mathbb{Z}^k$ -version of Furstenberg's correspondence principle (see, for example, Proposition 7.2 in [BMc]).

**Proposition 1.1.** *Given  $E \subset \mathbb{Z}^k$  with  $d^*(E) > 0$ , there is a probability space  $(X, \mathcal{B}, \mu)$ , commuting invertible measure preserving transformations  $T_1, T_2, \dots, T_k$  of  $X$  and  $A \in \mathcal{B}$  with  $d^*(E) = \mu(A)$  such that for any  $m \in \mathbb{N}$  and  $\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_m \in \mathbb{Z}^k$  one has*

$$d^*(E \cap (E - \mathbf{n}_1) \cap (E - \mathbf{n}_2) \cap \dots \cap (E - \mathbf{n}_m)) \geq \mu(A \cap T^{-\mathbf{n}_1} A \cap \dots \cap T^{-\mathbf{n}_m} A),$$

where for  $\mathbf{n} = (n_1, \dots, n_k)$ ,  $T^{\mathbf{n}} = T_1^{n_1} \dots T_k^{n_k}$ .

It follows from the result obtained in [BKMST] that both Theorems 1.1 and 1.2 remain true if one replaces  $[p_n^{c_i}]$ ,  $i = 1, 2, \dots, k$ , by  $[(p_n - h)^{c_i}]$  for any fixed  $h \in \mathbb{Z}$ . Now, when  $h = \pm 1$ , these results hold (in a slightly modified form) for  $c_i \in \mathbb{N}$  as well and one has the following theorem which provides a simultaneous extension of various classical results.  $\mathcal{P}$  denotes the set of prime numbers  $\{2, 3, 5, 7, \dots\}$ .

**Theorem 1.4** (Theorem 5.1 and Corollary 5.1 in [BKMST]). *Let*

$$D_h = \{((p - h)^{\alpha_1}, \dots, (p - h)^{\alpha_k}, [(p - h)^{\beta_1}], \dots, [(p - h)^{\beta_l}]) \mid p \in \mathcal{P}\},$$

where  $\alpha_1, \dots, \alpha_k$  are positive integers and  $\beta_1, \dots, \beta_l$  are positive non-integers.

- (i) *For any measure preserving  $\mathbb{Z}^{k+l}$ -action  $(T^{\mathbf{d}})_{\mathbf{d} \in \mathbb{Z}^{k+l}}$  on a probability space  $(X, \mathcal{B}, \mu)$ ,  $\{\mathbf{d} \in D_h : \mu(A \cap T^{-\mathbf{d}} A) > \mu^2(A) - \epsilon\}$  has positive lower relative density in  $D_h$  for  $h = \pm 1$ .<sup>1</sup>*
- (ii) *If  $E \subset \mathbb{Z}^{k+l}$  with  $d^*(E) > 0$ , then for any  $\epsilon > 0$*

$$\{\mathbf{d} \in D_h : d^*(E \cap E - \mathbf{d}) \geq d^*(E)^2 - \epsilon\}$$

*has positive lower relative density in  $D_h$  for  $h = \pm 1$ . Furthermore,*

$$\liminf_{N \rightarrow \infty} \frac{|\{p \leq N : ((p - 1)^{\alpha_1}, \dots, (p - 1)^{\alpha_k}, [(p - 1)^{\beta_1}], \dots, [(p - 1)^{\beta_l}]) \in E - E\}|}{\pi(N)} > 0.$$

$$\liminf_{N \rightarrow \infty} \frac{|\{p \leq N : ((p + 1)^{\alpha_1}, \dots, (p + 1)^{\alpha_k}, [(p + 1)^{\beta_1}], \dots, [(p + 1)^{\beta_l}]) \in E - E\}|}{\pi(N)} > 0.$$

Theorems 1.1 and 1.4 are derived in [BKMST] with the help of the following equidistribution result.

<sup>1</sup> For sets  $A \subset B \subset \mathbb{Z}^m$ , the relative density and the lower relative density of  $A$  with respect to  $B$  are defined as

$$\lim_{n \rightarrow \infty} \frac{|A \cap [-n, n]^m|}{|B \cap [-n, n]^m|} \quad \text{and} \quad \liminf_{n \rightarrow \infty} \frac{|A \cap [-n, n]^m|}{|B \cap [-n, n]^m|}.$$

**Theorem 1.5** (Theorem 2.1 in [BKMST]). *Let  $\xi(x) = \sum_{j=1}^m \alpha_j x^{\theta_j}$ , where  $0 < \theta_1 < \theta_2 < \dots < \theta_m$ ,  $\alpha_j$  are non-zero reals, and assume that if all  $\theta_j \in \mathbb{N}$ , then at least one  $\alpha_j$  is irrational. Then the sequence  $(\xi(p))_{p \in \mathcal{P}}$  is u.d. mod 1.<sup>2</sup>*

Note that in the case when all  $\theta_j \in \mathbb{N}$ , Theorem 1.5 reduces to the classical result of Rhin [Rh] which states that if  $f(x)$  is a polynomial with at least one coefficient other than the constant term irrational, then  $(f(p))_{p \in \mathcal{P}}$  is uniformly distributed mod 1. Incidentally, in Theorem 3.5 below we provide a new short proof of (a slight extension of) Rhin's theorem. While Theorem 1.5 forms a rather natural extension of Rhin's result, one would like to know whether the phenomenon of uniform distribution along primes holds for more general regularly behaving and, say, eventually monotone functions. Besides being of independent interest, any such extension of Theorem 1.5 allows one to obtain new applications to ergodic theory and combinatorics. In this context it is natural to consider functions belonging to Hardy fields.

Let  $B$  denote the set of germs at  $+\infty$  of continuous real functions on  $\mathbb{R}$ . Note that  $B$  forms a ring with respect to pointwise addition and multiplication.

**Definition 1.1.** A *Hardy field* is any subfield of  $B$  which is closed under differentiation. By  $\mathbf{U} \subset B$  we denote the union of all Hardy fields.

A classical example of a Hardy field is provided by field  $L$  of logarithmico-exponential functions introduced in [Har1, Har2], that is, the collection of all functions that can be constructed using the real constants, the functions  $e^x$  and  $\log x$  and the operations of addition, multiplication, division and composition of functions.

For any  $f \in \mathbf{U}$ ,  $\lim_{x \rightarrow \infty} f(x)$  exists as an element of  $\mathbb{R} \cup \{-\infty, \infty\}$ . This implies that periodic functions such as  $\sin x$  and  $\cos x$  do not belong to  $\mathbf{U}$ . Also if  $f_1$  and  $f_2$  belong to the same Hardy field, then the limit  $\lim_{x \rightarrow \infty} \frac{f_1(x)}{f_2(x)}$  exists (it may be infinite). See [Bos] and some references therein for more information about Hardy fields.

A function  $f \in \mathbf{U}$  is said to be subpolynomial if, for some  $n \in \mathbb{N}$ ,  $|f(x)| < x^n$  for all large enough  $x$ . It was proved in [Bos] that if  $f \in \mathbf{U}$  is a subpolynomial function, then  $(f(n))_{n \in \mathbb{N}}$  is uniformly distributed mod 1 if and only if for any  $P(x) \in \mathbb{Q}[x]$  one has  $\lim_{x \rightarrow \infty} \frac{f(x) - P(x)}{\log x} = \pm\infty$ . One of the main results of this paper states that an equidistribution result similar to Theorem 1.5 holds for any subpolynomial function satisfying Boshernitzan's condition.

**Theorem 1.6** (Theorem 3.1 in Section 3). *For a subpolynomial function  $f(x) \in \mathbf{U}$ , the following conditions are equivalent:*

- (1)  $(f(n))_{n \in \mathbb{N}}$  is u.d. mod 1.
- (2)  $(f(p))_{p \in \mathcal{P}}$  is u.d. mod 1.
- (3) For any  $P \in \mathbb{Q}[x]$ ,

$$\lim_{x \rightarrow \infty} \frac{f(x) - P(x)}{\log x} = \pm\infty.$$

We will give now a sample of various applications of Theorem 1.6 obtained in this paper. One of these applications is an extension of the above Theorem 1.1.

---

<sup>2</sup>We are tacitly assuming that the set  $\mathcal{P} = (p_n)_{n \in \mathbb{N}}$  is naturally ordered, so that  $(f(p))_{p \in \mathcal{P}}$  is just another way of writing  $(f(p_n))_{n \in \mathbb{N}}$ .

For a given Hardy field  $H$ , let  $\mathbf{H}$  be the set of all subpolynomial functions  $\xi \in H$  such that

$$\text{either } \lim_{x \rightarrow \infty} \frac{\xi(x)}{x^{l+1}} = \lim_{x \rightarrow \infty} \frac{x^l}{\xi(x)} = 0 \text{ for some } l \in \mathbb{N}, \text{ or } \lim_{x \rightarrow \infty} \frac{\xi(x)}{x} = \lim_{x \rightarrow \infty} \frac{\log x}{\xi(x)} = 0.$$

**Theorem 1.7** (Theorem 4.2 in Section 4). *Let  $U_1, \dots, U_k$  be commuting unitary operators on a Hilbert space  $\mathcal{H}$ . Let  $\xi_1, \dots, \xi_k \in \mathbf{H}$  such that  $\sum_{i=1}^k b_i \xi_i \in \mathbf{H}$  for any  $(b_1, \dots, b_k) \in \mathbb{R}^k \setminus \{(0, 0, \dots, 0)\}$ . Then,*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N U_1^{[\xi_1(p_n)]} \dots U_k^{[\xi_k(p_n)]} f = f^*,$$

strongly in  $\mathcal{H}$ , where  $f^*$  is the projection of  $f$  on  $\mathcal{H}_{inv} := \{f \in \mathcal{H} : U_i f = f \text{ for all } i\}$ .

**Theorem 1.8** (cf. Theorem 4.3 in Section 4). *Let  $\xi_1, \dots, \xi_k \in \mathbf{H}$  such that  $\sum_{i=1}^k b_i \xi_i \in \mathbf{H}$  for any  $(b_1, \dots, b_k) \in \mathbb{R}^k \setminus \{(0, 0, \dots, 0)\}$ . Let  $L : \mathbb{Z}^k \rightarrow \mathbb{Z}^m$  be a linear transformation and  $(\psi_1(n), \dots, \psi_m(n)) = L([\xi_1(n)], \dots, [\xi_k(n)])$ . Let  $T_1, T_2, \dots, T_m$  be commuting, invertible measure preserving transformations on a probability space  $(X, \mathcal{B}, \mu)$ . Then, for any  $A \in \mathcal{B}$  with  $\mu(A) > 0$ , one has*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mu(A \cap T_1^{-\psi_1(p_n)} \dots T_m^{-\psi_m(p_n)} A) \geq \mu^2(A).$$

Denote

$$\begin{aligned} \mathbf{D}_{-1} &= \{(P_1(p-1), \dots, P_l(p-1), [\xi_1(p)], \dots, [\xi_k(p)]) \mid p \in \mathcal{P}\}, \\ \mathbf{D}_1 &= \{(P_1(p+1), \dots, P_l(p+1), [\xi_1(p)], \dots, [\xi_k(p)]) \mid p \in \mathcal{P}\}, \end{aligned}$$

where  $P_1, \dots, P_l \in \mathbb{Z}[x]$  with  $P_i(0) = 0$  for all  $1 \leq i \leq l$  and  $\xi_1, \dots, \xi_k \in \mathbf{H}$  such that  $\sum_{j=1}^k b_j \xi_j(x) \in \mathbf{H}$  for any  $(b_1, \dots, b_k) \in \mathbb{Z}^k \setminus \{(0, 0, \dots, 0)\}$ .

**Theorem 1.9** (cf. Theorem 4.5 and Remark 4.2 in Section 4). *Enumerate the elements of  $\mathbf{D}_i$ , ( $i = \pm 1$ ), as follows:*

$$\mathbf{d}_{n,i} = (P_1(p_n + i), \dots, P_l(p_n + i), [\xi_1(p_n)], \dots, [\xi_k(p_n)]) \quad n = 1, 2, \dots$$

Let  $(T^{\mathbf{d}})_{\mathbf{d} \in \mathbb{Z}^{l+k}}$  be a measure preserving  $\mathbb{Z}^m$ -action on a probability space  $(X, \mathcal{B}, \mu)$ . Then  $\mathbf{D}_1$  and  $\mathbf{D}_{-1}$  are “averaging” sets of recurrence:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mu(A \cap T^{-\mathbf{d}_{n,i}} A) > 0 \quad (i = \pm 1).$$

Moreover, for any  $\epsilon > 0$ ,  $\{\mathbf{d} \in \mathbf{D}_i : \mu(A \cap T^{-\mathbf{d}} A) \geq \mu^2(A) - \epsilon\}$  has positive lower relative density in  $\mathbf{D}_i$  ( $i = \pm 1$ ).

**Theorem 1.10** (cf. Corollary 4.5 and Remark 4.2 in Section 4). *If  $E \subset \mathbb{Z}^{l+k}$  with  $d^*(E) > 0$ , then for any  $\epsilon > 0$ ,*

$$\{\mathbf{d} \in \mathbf{D}_i : d^*(E \cap E - \mathbf{d}) \geq d^*(E)^2 - \epsilon\}$$

has positive lower relative density in  $\mathbf{D}_i$  for  $i = \pm 1$ . Furthermore,

$$\liminf_{N \rightarrow \infty} \frac{|\{p \leq N : (P_1(p-1), \dots, P_l(p-1), [\xi_1(p)], \dots, [\xi_k(p)]) \in E - E\}|}{\pi(N)} > 0.$$

$$\liminf_{N \rightarrow \infty} \frac{|\{p \leq N : (P_1(p+1), \dots, P_l(p+1), [\xi_1(p)], \dots, [\xi_k(p)]) \in E - E\}|}{\pi(N)} > 0.$$

The structure of the paper is as follows. In Section 2 we establish various differential inequalities for functions in Hardy fields which are needed for the proofs of uniform distribution results in Section 3. We also collect in Section 2 various auxiliary number theoretical results. Section 4 is devoted to various applications. These include (some refinements of) the results formulated in this introduction as well as new results pertaining to sets of recurrence in  $\mathbb{Z}^d$ .

**Notation.** The following notation will be used throughout this paper.

- (1) We write  $e(x) = \exp(2\pi i x)$ .
- (2) For positive  $Y$ ,  $X \ll Y$  (or  $X = O(Y)$ ) means  $|X| \leq cY$  for some positive constant  $c$ .
- (3)  $X \asymp Y$  means  $c_1 X \leq Y \leq c_2 X$  for some positive constants  $c_1, c_2$ .
- (4)  $\sum_{p \leq N}$  denotes the sum over primes in the interval  $[1, N] = \{1, 2, \dots, N\}$ .
- (5)  $\phi(n)$  is Euler's totient function, which is defined as the number of positive integers  $\leq n$  that are relatively prime to  $n$ .
- (6) By  $\pi(x)$ , we denote the number of primes not exceeding  $x$ .
- (7) The Möbius function  $\mu : \mathbb{N} \rightarrow \mathbb{R}$  is defined by

$$\mu(n) = \begin{cases} 0 & \text{if } n \text{ has one or more repeated prime factors} \\ 1 & \text{if } n = 1 \\ (-1)^k & \text{if } n \text{ is a product of } k \text{ distinct primes.} \end{cases}$$

- (8) The von Mangoldt function  $\Lambda : \mathbb{N} \rightarrow \mathbb{R}$  is defined by

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^k \text{ for some prime } p \text{ and integer } k \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

- (9)  $\Lambda_1(n)$  denotes the characteristic function of the set of primes,

$$\Lambda_1(n) = \begin{cases} 1 & \text{if } n \text{ is a prime number} \\ 0 & \text{otherwise.} \end{cases}$$

The function  $\Lambda_1(n, r, a)$  is defined by

$$\Lambda_1(n, r, a) = \begin{cases} 1, & \text{if } n \text{ is a prime number and } n \equiv a \pmod{r} \\ 0, & \text{otherwise.} \end{cases}$$

- (10) For sets  $A \subset B \subset \mathbb{Z}^m$ , the relative density and the lower relative density of  $A$  with respect to  $B$  are defined as

$$\lim_{n \rightarrow \infty} \frac{|A \cap [-n, n]^m|}{|B \cap [-n, n]^m|} \quad \text{and} \quad \liminf_{n \rightarrow \infty} \frac{|A \cap [-n, n]^m|}{|B \cap [-n, n]^m|}.$$

- (11) We write  $f(x) \uparrow \infty$  if  $f(x)$  is eventually increasing and  $\lim_{x \rightarrow \infty} f(x) = \infty$  and  $f(x) \downarrow 0$  if  $f(x)$  is eventually decreasing and  $\lim_{x \rightarrow \infty} f(x) = 0$ .
- (12) By a slight abuse of notation we write  $Tf(x) = f(Tx)$  for a measure preserving transformation  $T$  and a function  $f$ .
- (13)  $\{x\}$  denotes the fractional part of a real number  $x$ .
- (14) The notation  $\|\cdot\|$  must be handled with care: in Sections 2 and 3, for  $x \in \mathbb{R}$ ,  $\|x\|$  denotes the distance to the nearest integer of  $x$  and in Section 4,  $\|f\|_{\mathcal{H}} = \langle f, f \rangle^{1/2}$  is the norm of  $f$  in a Hilbert space  $\mathcal{H}$ .

## 2. PRELIMINARIES

## 2.1. Differential inequalities for functions from Hardy fields.

Following Boshernitzan [Bos] we say that a subpolynomial function  $f(x) \in \mathbf{U}$  is of type  $x^{l+}$  ( $l = 0, 1, 2, \dots$ ) if

$$\lim_{x \rightarrow \infty} \frac{x^l}{f(x)} = \lim_{x \rightarrow \infty} \frac{f(x)}{x^{l+1}} = 0.$$

**Proposition 2.1** (cf. Ch. VI in [Har2]). *For any function  $f$  belonging to a Hardy field such that no derivative  $f^{(n)}$  satisfies  $|f^{(n)}| \asymp 1$ ,*

$$\left| \frac{xf^{(n+1)}(x)}{f^{(n)}(x)} \right| \gg \frac{1}{\log^2 x}. \quad (2.1)$$

*Proof.* Note that  $\lim_{x \rightarrow \infty} |f^{(n)}(x)|$  is either 0 or  $\infty$ . Let  $g(x) = f^{(n)}(x)$ . Suppose (2.1) does not hold, that is, for some  $C > 0$ ,

$$\left| \frac{xg'(x)}{g(x)} \right| \leq \frac{C}{\log^2 x}.$$

If  $\lim_{x \rightarrow \infty} |g(x)| = \infty$ , then  $\frac{g'(x)}{g(x)} \leq \frac{C}{x \log^2 x}$ . Integrating on both sides from  $a$  to  $\infty$ , where  $a$  is some positive constant, we obtain

$$\infty = \lim_{x \rightarrow \infty} \log |g(x)| - \log |g(a)| = \int_a^\infty \frac{C}{x \log^2 x} < \infty,$$

which gives a contradiction.

If  $\lim_{x \rightarrow \infty} |g(x)| = 0$ , then  $-\frac{g'(x)}{g(x)} \leq \frac{C}{x \log^2 x}$ . Again, integrating on both sides leads to a contradiction:

$$\infty = \lim_{x \rightarrow \infty} (-\log |g(x)|) + \log |g(a)| = \int_a^\infty \frac{C}{x \log^2 x} < \infty.$$

□

**Proposition 2.2.** *Let  $f(x) \in \mathbf{U}$  be a subpolynomial function. Suppose that*

$$f(x) \text{ is of the type } x^{0+} \text{ and } \lim_{x \rightarrow \infty} \left| \frac{f(x)}{\log x} \right| = \lim_{x \rightarrow \infty} x|f'(x)| = \infty. \quad (2.2)$$

*Then, for any  $j \geq 1$  and sufficiently large  $x$ , we have*

$$\frac{1}{2 \log x} \ll \frac{xf'(x)}{f(x)} \ll 1, \quad (2.3)$$

$$\frac{1}{\log^2 x} \ll \left| \frac{xf^{(j+1)}(x)}{f^{(j)}(x)} + j \right| \ll 1, \quad (2.4)$$

$$\frac{1}{\log^2 x} \ll \left| \frac{xf^{(j+1)}(x)}{f^{(j)}(x)} \right| \ll 1, \quad (2.5)$$

$$0 \ll \frac{f'(x)}{f'(2x)} \ll \log x. \quad (2.6)$$

*Proof.* Without loss of generality, we can assume that

$$f(x) \uparrow \infty, \quad f'(x) \downarrow 0, \quad xf'(x) \uparrow \infty.$$

We start by proving (2.3). Consider  $g(x) = f(x) - xf'(x)$ . Then  $g'(x) = -xf''(x)$ . Since  $f'(x) \downarrow 0$ ,  $f''(x) < 0$ , so  $g'(x) > 0$  eventually. Hence there exists  $K$  such that  $g(x) > K$ . Then  $\frac{xf'(x)}{f(x)} \leq 1 - \frac{K}{f(x)}$ . Since  $f(x) \uparrow \infty$ , we have

$$\frac{xf'(x)}{f(x)} \ll 1.$$

Now consider  $h(x) := xf'(x) \log x - f(x)$ . Then  $h'(x) = (xf''(x) + f'(x)) \log x$ . Since  $xf'(x)$  is increasing,  $xf''(x) + f'(x) = (xf'(x))' \geq 0$ . Thus,  $h'(x) \geq 0$ ,  $h(x) = xf'(x) \log x - f(x) > C$  for some constant  $C$ . Then

$$\frac{xf'(x)}{f(x)} \geq \left(1 + \frac{C}{f(x)}\right) \frac{1}{\log x}.$$

Since  $f(x) \uparrow \infty$ , (2.3) follows.

We turn now our attention to formulas (2.4) and (2.5). By L'Hospital's rule,

$$\lim_{x \rightarrow \infty} \frac{xf'(x)}{f(x)} = \lim_{x \rightarrow \infty} \frac{xf^{(j+1)}(x)}{f^{(j)}(x)} + j.$$

From (2.3),

$$\left| \frac{xf^{(j+1)}(x)}{f^{(j)}(x)} + j \right| \ll 1, \quad \left| \frac{xf^{(j+1)}(x)}{f^{(j)}(x)} \right| \ll 1.$$

From (2.1),

$$\frac{1}{\log^2 x} \ll \left| \frac{xf^{(j+1)}(x)}{f^{(j)}(x)} \right|.$$

Now let  $q(x) = x^j f^{(j)}(x)$ . Note that from  $\lim_{x \rightarrow \infty} xf'(x) = \infty$ , by L'Hospital's rule

$$\lim_{x \rightarrow \infty} |x^2 f''(x)| = \lim_{x \rightarrow \infty} \left| \frac{f'(x)}{1/x} \right| = \infty.$$

By a similar argument, we have  $\lim_{x \rightarrow \infty} |x^k f^{(k)}(x)| = \infty$  for all  $k$ . Now,  $\lim_{x \rightarrow \infty} f'(x) = 0$  implies  $\lim_{x \rightarrow \infty} |x^m f^{(n)}(x)| = 0$  for any non-negative integers  $m < n$ . Thus no  $q^{(n)}$  satisfy  $q^{(n)} \asymp 1$ . Then from (2.1),

$$\frac{1}{\log^2 x} \ll \left| \frac{xg'(x)}{g(x)} \right| = \left| \frac{xf^{(j+1)}(x)}{f^{(j)}(x)} + j \right|.$$

Finally, let us prove (2.6). From (2.3), we have

$$\frac{f(x)}{2x \log x} \ll f'(x) \ll \frac{f(x)}{x}.$$

Hence

$$\frac{f'(x)}{f'(2x)} \ll \frac{f(x)}{x} \cdot \frac{2(2x) \log 2x}{f(2x)} = 4 \frac{f(x)}{f(2x)} \log 2x \leq 4 \log 2x,$$

since  $f(x) \leq f(2x)$ . □

**Remark 2.1.** The proof actually shows that in the formulas (2.1), (2.4) and (2.5) one can replace  $\log^2 x$  with  $\log^{1+\epsilon} x$  for any  $\epsilon > 0$ .

**Proposition 2.3.** *Let  $f(x) \in \mathbf{U}$  be a subpolynomial function. If  $f(x)$  is of the type  $x^{l+}$  for some  $l \geq 1$ , then for any non-negative integer  $j$  and any  $\epsilon > 0$*

$$\left| j + \frac{xf^{(j+1)}(x)}{f^{(j)}(x)} \right| \asymp 1 \quad (2.7)$$

and

$$x^{\beta-j-\epsilon} \ll |f^{(j)}(x)| \ll x^{\beta-j+\epsilon} \quad (2.8)$$

for some fixed  $\beta \in [l, l+1]$ .

*Proof.* Without loss of generality we can assume that  $f(x) > 0$  eventually. First, by L'Hospital's rule,

$$\beta = \lim_{x \rightarrow \infty} \frac{\log f(x)}{\log x} = \lim_{x \rightarrow \infty} \frac{xf'(x)}{f(x)}. \quad (2.9)$$

Since

$$\lim_{x \rightarrow \infty} \frac{x^l}{f(x)} = \lim_{x \rightarrow \infty} \frac{l!}{f^{(l)}(x)} = 0 \quad \text{and} \quad \lim_{x \rightarrow \infty} \frac{f(x)}{x^{l+1}} = \lim_{x \rightarrow \infty} \frac{f^{(l+1)}(x)}{(l+1)!} = 0,$$

we have  $x^l \ll f(x) \ll x^{l+1}$ , which proves  $l \leq \beta \leq l+1$ .

Again, by L'Hospital's rule, if  $\lim_{x \rightarrow \infty} \frac{(xf'(x))^{(j)}}{(f(x))^{(j)}} = \beta$ , then  $\lim_{x \rightarrow \infty} \frac{(xf'(x))^{(j+1)}}{(f(x))^{(j+1)}} = \beta$  since

- (1)  $\lim_{x \rightarrow \infty} f^{(j)}(x) = \pm\infty$  or 0 from that  $f(x)$  is of the type  $x^{l+}$ .
- (2)  $\lim_{x \rightarrow \infty} f^{(j)}(x) = \lim_{x \rightarrow \infty} (xf'(x))^{(j)} = \pm\infty$  or 0 from that  $0 < \beta < \infty$ .

Note now that

$$\frac{(xf'(x))^{(j)}}{f^{(j)}(x)} = j + \frac{xf^{(j+1)}(x)}{f^{(j)}(x)},$$

which proves (2.7). Also, (2.9) implies that for any  $\epsilon > 0$  we have

$$x^{\beta-\epsilon} \ll f(x) \ll x^{\beta+\epsilon}.$$

Now we claim that for any  $a \neq 0$ ,

$$\lim_{x \rightarrow \infty} \frac{f^{(j)}(x)}{x^a} = \lim_{x \rightarrow \infty} \frac{f^{(j+1)}(x)}{ax^{a-1}}. \quad (2.10)$$

Choose  $b \in \mathbb{R}$  such that both  $f^{(j)}(x)x^b$  and  $x^{a+b}$  converge to 0 as  $x \rightarrow \infty$ . Then by applying L'Hospital's rule we get

$$\lim_{x \rightarrow \infty} \frac{f^{(j)}(x)}{x^a} = \lim_{x \rightarrow \infty} \frac{x^b f^{(j)}(x)}{x^{a+b}} = \lim_{x \rightarrow \infty} \frac{b}{a+b} \frac{f^{(j)}(x)}{x^a} + \lim_{x \rightarrow \infty} \frac{1}{a+b} \frac{f^{(j+1)}(x)}{x^{a-1}}.$$

Thus,

$$\frac{a}{a+b} \lim_{x \rightarrow \infty} \frac{f^{(j)}(x)}{x^a} = \frac{1}{a+b} \frac{f^{(j+1)}(x)}{x^{a-1}}.$$

Dividing both sides with  $\frac{a}{a+b}$ , we obtain (2.10). Hence,

$$x^{\beta-j-\epsilon} \ll |f^{(j)}(x)| \ll x^{\beta-j+\epsilon},$$

which proves (2.8). □



### 2.2. Some classical number-theoretic lemmas.

We collect in this subsection some classical results which will be needed for the proofs in the next section. Recall that  $\pi(x)$  is the number of primes less than or equal to  $x$ .

**Lemma 2.1** (Erdős-Turán; cf. Theorem 1.21 in [DT] or Theorem 2.5 in [KN]). *For a sequence  $(a_n)_{n \in \mathbb{M}}$  of real numbers, we define*

$$D(x) := \sup_{[\alpha, \beta] \subset [0, 1]} \left| \frac{1}{\pi(x)} |\{p \in [1, x] : \{a_p\} \in [\alpha, \beta]\}| - (\beta - \alpha) \right|.$$

*For any  $Q$  we have*

$$D(x) \ll \frac{1}{Q} + \frac{1}{\pi(x)} \sum_{q \leq Q} \frac{1}{q} \left| \sum_{p \leq x} e(qa_p) \right|.$$

**Lemma 2.2.** *Let  $(a_n)_{n \in \mathbb{N}}$  be a sequence of real numbers. Suppose that for any large enough  $X$ , there exists  $Q = Q(X) \uparrow \infty$  such that for all  $1 \leq q \leq Q$ ,*

$$S := \sum_{X_0 < p \leq X} e(qa_p) \ll \frac{\pi(X)}{Q}, \quad (2.11)$$

*where  $X_0 = \frac{X}{Q}$ . Then  $(a_p)_{p \in \mathcal{P}}$  is u.d. mod 1.*

*Proof.* Let

$$D(X) := \sup_{[\alpha, \beta] \subset [0, 1]} \left| \frac{1}{\pi(X)} |\{p \in [1, X] : \{a_p\} \in [\alpha, \beta]\}| - (\beta - \alpha) \right|.$$

Let  $Q_1 = Q_1(X) = \frac{Q}{\log Q}$ . Then  $\lim_{X \rightarrow \infty} Q_1(X) = \infty$ . From (2.11),

$$\frac{1}{\pi(X)} \sum_{q \leq Q} \frac{1}{q} \left| \sum_{X_0 < p \leq X} e(qa_p) \right| \ll \frac{1}{Q_1}. \quad (2.12)$$

Since

$$\frac{1}{\pi(X)} \sum_{q \leq Q} \frac{1}{q} \left| \sum_{p \leq X_0} e(qa_p) \right| \ll \frac{1}{\pi(X)} \sum_{q \leq Q} \frac{1}{q} \pi\left(\frac{X}{Q}\right) \ll \frac{\log Q}{Q},$$

we have

$$\frac{1}{\pi(X)} \sum_{q \leq Q} \frac{1}{q} \left| \sum_{p \leq X} e(qa_p) \right| \ll \frac{1}{Q_1}. \quad (2.13)$$

By Erdős-Turán inequality (Lemma 2.1),

$$D(X) \ll \frac{1}{Q} + \frac{1}{Q_1} \rightarrow 0 \quad \text{as } X \rightarrow \infty. \quad \square$$

**Lemma 2.3** (Prime number theorem).

(1) (cf. [Kar, Theorem 2 in Ch. V])

$$\pi(x) := \sum_{p \leq x} 1 = Li(x) + R(x),$$

where  $Li(x) = \int_2^x \frac{dt}{\log t}$  and  $R(x) \ll x \exp(-C\sqrt{\log x})$  for some positive constant  $C$ .

(2) (cf. [Kar, Theorem 6 in Ch. IX]) Let  $(a, q) = 1$ . By  $\pi(x; q, a)$  we denote the number of primes less than or equal to  $x$  that are congruent to  $a \pmod q$ . Then, for some positive constant  $C$ ,

$$\pi(x; q, a) = \frac{1}{\phi(q)} Li(x) - E \frac{\chi(a)}{\phi(q)} \int_2^x \frac{u^{\beta_1-1}}{\log u} du + O(x \exp(-C\sqrt{\log x})),$$

where  $E = 1$  if there exists a real character  $\chi$  modulo  $q$  such that  $L(s, \chi)$  has a real zero  $\beta_1$  in  $(1-c/\log q, 1)$  (so-called Siegel zero) and  $E = 0$  otherwise.

The following lemma is a corollary of the previous result.

**Lemma 2.4** (Siegel - Walfisz, cf. Corollary 2 in Ch. IX [Kar]). For every  $A > 0$ , there exists  $C_A > 0$  such that for any  $x$ ,  $2 \leq q \leq (\log x)^A$  and  $(a, q) = 1$ ,

$$\pi(x; q, a) = \frac{1}{\phi(q)} Li(x) + O(x \exp(-C_A \sqrt{\log x})).$$

**Lemma 2.5** (Partial summation formula). Let  $(a_n)_{n \in \mathbb{N}}$  and  $(b_n)_{n \in \mathbb{N}}$  be two sequences. For any non-negative integers  $X_1, X_2$  with  $X_1 < X_2$

$$\sum_{n=X_1}^{X_2} a_n b_n = \sum_{n=X_1}^{X_2-1} \left( (a_n - a_{n+1}) \sum_{m=X_1}^n b_m \right) + a_{X_2} \sum_{m=X_1}^{X_2} b_m.$$

Also, let  $\sum_{m=X_1}^x b_m = B(x) + R(x)$  with  $|R(x)| \leq R$  for  $X_1 - 1 \leq x \leq X_2$  and let  $|a_n| \leq A$ . Then

$$\sum_{n=X_1}^{X_2} a_n b_n = \sum_{n=X_1}^{X_2} a_n (B(n) - B(n-1)) + O(RA + R \sum_{n=X_1}^{X_2-1} |a_n - a_{n+1}|).$$

*Proof.* The first part is a well-known partial summation formula. To prove the second part, we write

$$\begin{aligned} \sum_{n=X_1}^{X_2} a_n b_n &= \sum_{n=X_1}^{X_2} a_n (B(n) + R(n) - B(n-1) - R(n-1)) \\ &= \sum_{n=X_1}^{X_2} a_n (B(n) - B(n-1)) + \sum_{n=X_1}^{X_2} a_n (R(n) - R(n-1)). \end{aligned}$$

Then the second sum can be estimated as follows:

$$\begin{aligned} \left| \sum_{n=X_1}^{X_2} a_n (R(n) - R(n-1)) \right| &= \left| \sum_{n=X_1}^{X_2-1} (a_n - a_{n+1}) R(n) + a_{X_2} R(X_2) - a_{X_1} R(X_1 - 1) \right| \\ &\leq 2AR + R \sum_{n=X_1}^{X_2-1} |a_n - a_{n+1}|. \end{aligned}$$

This completes the proof of the second part of the lemma.  $\square$

**Lemma 2.6** (Weyl - van der Corput; see Lemma 2.5 in [GK]). Suppose that  $(\xi(n))_{n \in \mathbb{N}}$  is a complex valued sequence. For any positive integer  $H$  and any interval

$I$  in  $\mathbb{N}$  of the form  $(a, b] = \{a+1, a+2, \dots, b\}$ , we have

$$\left| \sum_{n \in I} \xi(n) \right|^2 \leq \frac{|I|+H}{H} \sum_{|h| \leq H} \left( 1 - \frac{|h|}{H} \right) \sum_{n, n+h \in I} \xi(n) \overline{\xi(n+h)}.$$

**Lemma 2.7** ([Va]). *For any positive integers  $u, v$  and any  $X$  we have*

$$\sum_{v \leq n \leq X} \Lambda(n) g(n) = T_1 - T_2 - T_3,$$

where

$$\begin{aligned} T_1 &= \sum_{d \leq u} \sum_{m \leq X/d} \mu(d) (\log m) g(dm), \\ T_2 &= \sum_{m \leq uv} \sum_{r \leq X/m} a(m) g(mr), \quad a(m) = \sum_{d \leq u} \sum_{\substack{n \leq v \\ dn=m}} \mu(d) \Lambda(n) \\ T_3 &= \sum_{m > u} \sum_{v < n \leq X/m} b(m) \Lambda(n) g(mn), \quad b(m) = \sum_{\substack{d \leq u \\ d|m}} \mu(d). \end{aligned}$$

**Remark 2.2.** For  $a(m)$  and  $b(m)$  as defined in Lemma 2.7, the following estimates will be useful:

- $\sum_{m \leq uv} \frac{|a(m)|}{m} = \sum_{m \leq uv} \frac{1}{m} \sum_{d \leq u} \sum_{\substack{n \leq v \\ dn=m}} \mu(d) \Lambda(n) \ll \sum_{d \leq u} \frac{|\mu(d)|}{d} \sum_{n \leq v} \frac{\Lambda(n)}{n} \ll \log u \cdot \log^2 v.$
- $\sum_{y \leq m \leq 2y} |b(m)|^2 = \sum_{y \leq m \leq 2y} \left| \sum_{\substack{d \leq u \\ d|m}} \mu(d) \right|^2 = \sum_{d_1, d_2 \leq u} \mu(d_1) \mu(d_2) \sum_{\substack{y \leq m \leq 2y \\ d_1 | m, d_2 | m}} 1$   
 $= \sum_{d \leq u} \sum_{d_1 \leq \frac{u}{d}} \sum_{d_2 \leq \frac{u}{d}} \mu(dd_1 d_2) \frac{2y - y}{dd_1 d_2} \leq y \sum_{d, d_1, d_2 \leq u} \frac{1}{dd_1 d_2} \ll y \log^3 u.$
- $\sum_{y \leq m, m+h \leq 2y} |b(m)| |b(m+h)| \ll \sum_m (b^2(m) + b^2(m+h)) \ll y \log^3 u.$
- $\sum_{y \leq n \leq 2y} \Lambda^2(n) \ll \sum_{y \leq p \leq 2y} \log^2 p + \sqrt{y} \log^2 y \ll y \log y.$
- $\sum_{y \leq n, n+h \leq 2y} \Lambda(n) \Lambda(n+h) \leq \frac{1}{2} \sum_n (\Lambda^2(n) + \Lambda^2(n+h)) \ll y \log y.$

For the estimate of  $\sum \Lambda^2(n)$ , we used the following observation

$$|\{p^k : y \leq p^k \leq 2y, k \geq 2\}| \leq |\{n \in \mathbb{N} : n^k \leq 2y, k \geq 2\}| \leq \sqrt{y} \log y$$

and the fact that  $\Lambda(p^k) \leq \log 2y$  if  $p^k \leq 2y$ .

### 2.3. Some auxiliary results regarding exponential sums.

**Lemma 2.8** (cf. p.34 in [Mo]).

$$\frac{1}{q} \sum_{j=1}^q e \left( \frac{(n-b)j}{q} \right) = \begin{cases} 1 & n \equiv b \pmod{q} \\ 0 & \text{otherwise} \end{cases} \quad (2.14)$$

for any  $q \in \mathbb{N}$  and  $b \in \mathbb{N}$  with  $1 \leq b \leq q$ .

In the following lemma,  $I$  denotes an interval in  $\mathbb{N}$  of the form  $(a, b] = \{a+1, a+2, \dots, b\}$ , where  $(a, b \in \mathbb{N})$ . Note that in the formulation of Lemma 2.9  $\|\cdot\|$  denotes the distance to the closest integer.

**Lemma 2.9** (Kusmin - Landau; see Theorem 2.1 in [GK]). *If  $f(x)$  is continuously differentiable,  $f'(x)$  is monotonic, and  $0 < \lambda \leq \|f'(x)\| \leq \frac{1}{2}$  on some interval  $I$ , then*

$$\left| \sum_{n \in I} e(f(n)) \right| \ll \frac{1}{\lambda}.$$

The following lemma is an improvement of Lemma 2.5 in [BKMST].

**Lemma 2.10.** *Let  $k \geq 1$  and denote  $K = 2^k$ . Let  $f(x)$  be a  $(k+1)$ -times continuously differentiable real function on  $I = (X_1, X_1 + X] \subset (X_1, 2X_1]$ , where  $X, X_1 \in \mathbb{N}$ . Assume that  $f^{(k+1)}(x)$  is monotone on  $I$  and  $\lambda \leq |f^{(k+1)}(x)| \leq \alpha\lambda$  for some  $\lambda, \alpha > 0$ . Then we have*

$$\begin{aligned} S &:= \left| \sum_{x \in I} e(f(x)) \right| \\ &\ll X \left[ (\alpha\lambda)^{1/(2K-2)} + (\lambda X^{k+1})^{-1/K} (\log X)^{k/K} + \left( \frac{\alpha \log^k X}{X} \right)^{1/K} \right], \end{aligned} \quad (2.15)$$

where the implied constant depends on  $k$  only.

To prove Lemma 2.10 we need the following result, which follows from Lemma 2.6 by iteration.

**Lemma 2.11** (cf. [GK, Lemma 2.7]). *Let  $k$  be a positive integer and  $K = 2^k$ . Assume that  $I = (X_1, X_1 + X] \subset (X_1, 2X_1]$  and let  $S = \left| \sum_{x \in I} e(f(x)) \right|$ . For any positive  $H_1, \dots, H_k \ll X$ , we have*

$$\begin{aligned} \left( \frac{S}{X} \right)^K &\leq 8^{K-1} \left\{ \frac{1}{H_1^{K/2}} + \frac{1}{H_2^{K/4}} + \dots + \frac{1}{H_k} + \right. \\ &\quad \left. \frac{1}{H_1 \dots H_k X} \sum_{h_1=1}^{H_1} \dots \sum_{h_k=1}^{H_k} \left| \sum_{x \in I(\underline{h})} e(f_1(x)) \right| \right\}, \end{aligned} \quad (2.16)$$

where  $f_1(x) := f(\underline{h}, x) = h_1 \dots h_k \int_0^1 \dots \int_0^1 \frac{\partial^k}{\partial x^k} f(x + \underline{h} \cdot \underline{t}) d\underline{t}$ ,  $\underline{h} = (h_1, \dots, h_k)$ ,  $\underline{t} = (t_1, \dots, t_k)$  and  $I(\underline{h}) = (X_1, X_1 + X - h_1 - \dots - h_k]$ .

*Proof of Lemma 2.10.* Using first (2.16) and then Lemma 2.9 we obtain

$$\begin{aligned} |S/X|^K &\ll (1/H_1)^{(K/2)} + \dots + 1/H_k + 1/(H_1 \dots H_k X) \sum_{h_1=1}^{H_1} \dots \sum_{h_k=1}^{H_k} \frac{1}{\lambda h_1 \dots h_k} \\ &\ll (1/H_1)^{(K/2)} + \dots + 1/H_k + (\log X)^k / (\lambda X H_1 \dots H_k). \end{aligned} \quad (2.17)$$

Now we need to minimize the last expression subject to the conditions

$$H_j \leq X/(2k) \quad (j = 1, \dots, k) \quad \text{and} \quad 2\alpha\lambda H_1 \dots H_k \leq 1. \quad (2.18)$$

If  $\alpha\lambda X^k \leq 1$ , then we take  $H_j = X/(2k)$  and get

$$|S/X|^K \ll 1/X + (\log X)^k/(\lambda X^{(k+1)}),$$

which implies (2.15). Assume now that  $\alpha\lambda X^k > 1$ . If  $2\alpha\lambda X^{(2K-2)/K} \geq 1$ , that is, if  $1/X \leq (2\alpha\lambda)^{K/(2K-2)}$ , then we take  $H_1 = (1/(2\alpha\lambda))^{1/(K-1)} \leq X^{2/K}$  and  $H_j = (H_1)^{2^{j-1}}$  ( $j = 2, \dots, k$ ). Note that  $2\alpha\lambda H_1 H_2 \cdots H_k = 2\alpha\lambda H_1^{1+2+2^2+\dots+K/2} = 1$ , so that (2.18) is satisfied and the last term in (2.17) is  $2\alpha \frac{\log^k X}{X}$ . Hence, we obtain

$$|S/X|^K \ll (\alpha\lambda)^{K/(2K-2)} + \alpha \log^k X/X,$$

which proves (2.15) in case  $\alpha\lambda X^k > 1$  and  $2\alpha\lambda X^{(2K-2)/K} \geq 1$ .

To complete the proof of (2.15), we need to establish the desired estimate for the case  $1/X^k \leq 2\alpha\lambda \leq X^{(2-2K)/K}$ . Assume that, for some  $j = 1, 2, \dots, k-1$ ,

$$X^{\frac{2J}{K}-j-2} \leq 2\alpha\lambda \leq X^{\frac{J}{K}-j-1}, \quad (2.19)$$

where  $J = 2^j$ . This covers all the remaining possibilities. Take  $H_k = \cdots = H_{k-j+1} = X/(2k)$ ,  $H_2 = (H_1)^2$ ,  $H_3 = (H_1)^4$ ,  $H_{k-j} = (H_1)^{K/(2J)}$  so that

$$2\alpha\lambda X^j H_1 H_2 \cdots H_{k-j} = 2\alpha\lambda X^j H_1^{1+2+\dots+K/(2J)} = 1,$$

so that  $H_1 = (2\alpha\lambda X^j)^{J/(J-K)} \leq X^{2J/K}$ . Then  $2\alpha\lambda H_1 \cdots H_k = (1/2k)^j$ , and the last term in (2.17) is  $2\alpha \frac{1}{(2k)^j} \frac{\log^k X}{X}$ . Thus (2.17) will acquire the following form:

$$|S/X|^K \ll (\alpha\lambda X^j)^{KJ/(2K-2J)} + \alpha \log^k X/X.$$

We will show now with the help of (2.19) that  $(\alpha\lambda X^j)^{J/(2K-2J)} \ll (\alpha\lambda)^{1/(2K-2)}$ . Indeed we have

$$\begin{aligned} (\alpha\lambda X^j)^{J/(2K-2J)} &\ll (\alpha\lambda)^{1/(2K-2)} \Leftrightarrow (\alpha\lambda X^j)^{J(2K-2)} \ll (\alpha\lambda)^{2K-2J} \\ &\Leftrightarrow (\alpha\lambda)^{2KJ+2J} X^{jJ(2K-2)} \ll 1 \\ &\Leftrightarrow \alpha\lambda X^{j(K-1)/(K+1)} \ll 1. \end{aligned}$$

Since

$$\alpha\lambda X^{j(K-1)/(K+1)} \leq \alpha\lambda X^j \leq X^{J/K-1} \leq 1,$$

the proof of the lemma is completed.  $\square$

**Lemma 2.12** (Theorem 1, p.133 [Vin]). *Let  $f(x) = \alpha_1 x + \cdots + \alpha_t x^t$ , where one of the  $\alpha_i$  ( $1 \leq i \leq t$ ) is irrational. For a given positive integer  $X$ , write all coefficients of  $f(x)$  in the form*

$$\begin{aligned} \alpha_j &= \frac{e_j}{q_j} + z_j \quad \text{with} \quad (e_j, q_j) = 1, \quad 0 < q_j \leq \tau_j := X^{j/2} \\ |z_j| &\leq \frac{1}{q_j \tau_j}, \quad z_j = \delta_j X^{-j} \quad (j = 1, 2, \dots, t). \end{aligned} \quad (2.20)$$

Let

$$\delta_0 = \max\{|\delta_1|, \dots, |\delta_t|\} \quad \text{and} \quad \mathcal{Q} = [q_1, \dots, q_t]. \quad (2.21)$$

Denote

$$\begin{aligned} \nu &= \frac{1}{t}, \quad \rho = \frac{1}{17t^2(2 \log t + \log \log t + 2.8)}, \quad t \geq 12, \\ H &= \exp\left(\frac{(\log \log X)^2}{\log(1 + \epsilon_0)}\right) \end{aligned}$$

for some small positive  $\epsilon_0$ .

If  $\mathcal{Q} \leq X^{\nu/5}$  and  $\delta_0 \leq X^\nu$ , put  $\Delta_1 = (m, \mathcal{Q})^{\nu/2} \mathcal{Q}^{-\nu/2+\epsilon_0}$  if  $\delta_0 < 1$  and  $\Delta_1 = \delta_0^{-\nu/2} \mathcal{Q}^{-\nu/2+\epsilon_0}$  if  $\delta_0 \geq 1$ ; otherwise put  $\Delta_1 = X^{-\rho}$ .

Assuming  $m \leq (\Delta_1)^{-2}$ , we have

$$\left| \sum_{p \leq X} e(mf(p)) \right| \ll HX\Delta_1.$$

### 3. UNIFORM DISTRIBUTION

In this section we prove the main theorem of this paper:

**Theorem 3.1.** *For a subpolynomial function  $f \in \mathbf{U}$ , the following conditions are equivalent:*

- (1)  $(f(n))_{n \in \mathbb{N}}$  is u.d. mod 1.
- (2)  $(f(p))_{p \in \mathcal{P}}$  is u.d. mod 1.
- (3) For any  $P(x) \in \mathbb{Q}[x]$ ,

$$\lim_{x \rightarrow \infty} \frac{f(x) - P(x)}{\log x} = \pm \infty. \quad (3.1)$$

As was already mentioned in the introduction, the equivalence of conditions (1) and (3) was established in [Bos]. So it is sufficient to show  $(2) \Leftrightarrow (3)$ . The proof of the implication  $(2) \Rightarrow (3)$  is relatively routine and is given at the end of this section. The main task is to show  $(3) \Rightarrow (2)$ . In order to do this, we will use the classical method of reducing the evaluation of the sums  $\sum_{X_0 < p \leq X} e(qf(p))$ ,

where  $f(x) \in \mathbf{U}$  and  $q \in \mathbb{Z} \setminus \{0\}$ , to the estimation of expressions of the form  $\sum \Lambda(n)e(qf(n))$ . Indeed, it follows from Lemma 2.5 that

$$\begin{aligned} \sum_{X_0 < p \leq X} e(qf(p)) &= \sum_{X_0 < n \leq X} \frac{1}{\log n} \Lambda(n) e(qf(n)) + O(\sqrt{X}) \\ &\ll \frac{1}{\log X} \max_{X_0 < X_1 \leq X} \left| \sum_{X_0 \leq n \leq X_1} \Lambda(n) e(qf(n)) \right| + \sqrt{X}. \end{aligned}$$

The task of estimating the sums  $\sum \Lambda(n)e(qf(n))$  can, in turn, be reduced to estimating the classical expressions  $\sum e(qf(n))$ . In general, a subpolynomial function  $f(x) \in \mathbf{U}$  can be written as  $f(x) = f_1(x) + f_2(x)$ , where  $f_1(x)$  is of type  $x^{l+}$  for some  $l \geq 0$  and  $f_2(x) \in \mathbb{R}[x]$ . If  $f_1(x)$  is of type  $x^{l+}$  for some  $l \geq 1$  or  $f_1(x)$  is of type  $x^{0+}$  with an appropriate growth rate (this, in our context, means roughly that for some  $c > 0$  and all large  $x$ ,  $|f_1'(x)| \geq \frac{\log^c x}{x}$ ), then Lemmas 2.9 and 2.10 provide a good estimate for  $\sum_n e(qf(n))$ . Then Vaughan's identity (Lemma 2.7) and the estimates listed in Remark 2.2 allow one to get the desired estimate for  $\sum \Lambda(n)e(qf(n))$ . On the other hand, if  $f_1(x)$  is of "slow growth" one needs to employ other methods - the prime number theorem or Vinogradov's lemma (Lemma 2.12) - to estimate  $\sum_{p \leq X} e(qf(p))$ .

Accordingly, the proof of Theorem 3.1 will be achieved by separately considering for our function  $f = f_1 + f_2$  the following four cases:

- (a)  $f_1(x)$  is of the type  $x^{0+}$  and  $\lim_{x \rightarrow \infty} \frac{f_1(x)}{\log x} = \pm \infty$  and  $f_2(x) = 0$  (Theorem 3.2).
- (b)  $f_1(x)$  is of the type  $x^{l+}$  for some  $l \geq 1$  and  $f_2(x) = 0$  (Theorem 3.3).

- (c)  $f_1(x)$  is as in (a) and (b) and  $f_2(x) \in \mathbb{R}[x]$  (Theorem 3.4).  
 (d)  $f_1(x)$  is of the type  $x^{0+}$  and  $\lim_{x \rightarrow \infty} \left| \frac{f_1(x)}{\log x} \right| < \infty$  and  $f_2(x) \in \mathbb{R}[x]$  such that one of coefficients of  $f_2(x) - f_2(0)$  is irrational (Theorem 3.5).

**Theorem 3.2.** *Let  $f(x) \in \mathbf{U}$  be a subpolynomial function. Suppose that*

$$f(x) \text{ is of the type } x^{0+} \quad \text{and} \quad \lim_{x \rightarrow \infty} \left| \frac{f(x)}{\log x} \right| = \lim_{x \rightarrow \infty} x|f'(x)| = \infty. \quad (3.2)$$

*Then  $(f(p))_{p \in \mathcal{P}}$  is u.d. mod 1.*

*Proof.* In light of Lemma 2.2, it is sufficient to show that if  $Q(X) \leq \log X$  is an unbounded positive increasing function with

$$Q \leq \frac{1}{2} \left| f' \left( \frac{X}{\log X} \right) \right|^{-1} \quad \text{and} \quad Q \leq |X f'(X)|, \quad (3.3)$$

then

$$S := \sum_{X_0 < p \leq X} e(qf(p)) = \sum_{X_0 < n \leq X} \Lambda_1(n) e(qf(n)) \ll \frac{\pi(X)}{Q},$$

where  $X_0 = \frac{X}{Q}$  and  $\Lambda_1(n)$  is the characteristic function of the set of prime numbers.

Without loss of generality, we assume that  $f(x) \geq 0$  eventually, so  $f(x) \uparrow \infty$ ,  $f'(x) \downarrow 0$  and  $xf'(x) \uparrow \infty$ . We split the sum  $S$  as follows

$$S = S_0 + S_1 = \sum_{X_1 < p \leq X} e(qf(p)) + \sum_{X_0 < p \leq X_1} e(qf(p)),$$

where  $f'(n) \leq \frac{\log^{15} X}{X}$  for  $n > X_1$  and  $f'(n) > \frac{\log^{15} X}{X}$  for  $n \leq X_1$ .

Using Lemma 2.5 with  $a_n = e(qf(n))$  and  $b_n = \Lambda_1(n)$  we obtain (upon invoking Lemma 2.3)

$$\begin{aligned} S_0 &:= \sum_{X_1 < p \leq X} e(qf(p)) = \sum_{X_1 < n \leq X} \Lambda_1(n) e(qf(n)) \\ &= \sum_{n=X_1}^{X-1} (e(qf(n)) - e(qf(n+1))) (\pi(n) - \pi(X_1)) + e(qf(X)) (\pi(X) - \pi(X_1)) \\ &= \sum_{n=X_1}^{X-1} [e(qf(n)) - e(qf(n+1))] [Li(n) - Li(X_1)] + e(qf(X)) [Li(X) - Li(X_1)] \\ &\quad + O \left( \left( X \exp(-C\sqrt{\log X}) \right) \left( \sum_{n=X_1}^{X-1} |e(qf(n)) - e(qf(n+1))| + 1 \right) \right). \end{aligned} \quad (3.4)$$

Since

$$|e(qf(n)) - e(qf(n+1))| = |1 - e(q(f(n+1) - f(n)))| \ll q|f'(n)| \leq \frac{q \log^{15} X}{X},$$

the  $O$ -term is

$$\ll X \exp(-C\sqrt{\log X}) \left( \sum_{n \leq X} \frac{q \log^{15} X}{X} + 1 \right) \ll \frac{\pi(X)}{\log X}.$$

Now we use Lemma 2.5 with  $a_n = Li(n) - Li(X_1)$  and  $b_n = e(qf(n)) - e(qf(n+1))$  and obtain

$$\begin{aligned} S_0 &= \sum_{n=X_1}^X \int_n^{n+1} \frac{dx}{\log x} e(qf(n)) + O\left(\frac{\pi(X)}{\log X}\right) \\ &= \sum_{n=X_1}^X \frac{1}{\log n} e(qf(n)) + O\left(\frac{\pi(X)}{\log X}\right). \end{aligned} \quad (3.5)$$

From (3.3),  $\left|\frac{d}{dn}(qf(n))\right| = |qf'(n)| \leq \frac{1}{2}^3$ , so  $\|qf'\| > qf'(X)$  for  $Y \leq n \leq X$ . Use Lemma 2.5 with  $a_n = \frac{1}{\log n}$  and  $b_n = e(qf(n))$  and then use Lemma 2.9 and (3.3):

$$S_0 \ll \frac{1}{\log X} \frac{1}{qf'(X)} + \frac{\pi(X)}{Q} \ll \frac{\pi(X)}{Q}. \quad (3.6)$$

Now we need to evaluate

$$S_1 := \sum_{X_0 < p \leq X_1} e(qf(p)) = \sum_{X_0 < n \leq X_1} \frac{\Lambda(n)}{\log n} e(qf(n)) + O(\sqrt{X}).$$

To complete the proof we have to show that  $S_1 \ll \frac{\pi(X)}{Q}$ . Using Lemma 2.5 with  $a_n = \frac{1}{\log n}$  and  $b_n = \Lambda(n)e(qf(n))$  we obtain

$$|S_1| \ll \frac{1}{\log X} \max_{X_2 \leq X_1} \left| \sum_{X_0 < n \leq X_2} \Lambda(n)e(qf(n)) \right| + \sqrt{X}.$$

To estimate  $\sum_{X_0 < n \leq X_2} \Lambda(n)e(qf(n))$ , we divide the interval  $(X_0, X_2]$  into  $\ll \log Q$  subintervals of the form  $(X_3, X_4] \subset (X_3, 2X_3]$  and evaluate

$$S_2 := \sum_{X_3 < n \leq X_4} \Lambda(n)e(qf(n)).$$

Using Lemma 2.7 with  $g(n) = e(qf(n)) \cdot 1_{(X_3, X_4]}(n)$  and some  $u, v$  to be defined later, we obtain

$$|S_2| \leq |S_3| + |S_4| + |S_5| \quad (3.7)$$

where  $S_3, S_4$ , and  $S_5$  correspond to the sums  $T_1, T_2$  and  $T_3$  in Lemma 2.7:

$$\begin{aligned} S_3 &= \sum_{d \leq u} \sum_{\frac{X_3}{d} \leq m \leq \frac{X_4}{d}} \mu(d) (\log m) e(qf(dm)), \\ S_4 &= \sum_{m \leq uv} \sum_{\frac{X_3}{m} \leq r \leq \frac{X_4}{m}} a(m) e(qf(mr)), \quad a(m) = \sum_{d \leq u} \sum_{\substack{n \leq v \\ dn=m}} \mu(d) \Lambda(n) \\ S_5 &= \sum_{m > u} \sum_{\substack{\frac{X_3}{m} \leq n \leq \frac{X_4}{m} \\ v > n}} b(m) \Lambda(n) e(qf(mn)), \quad b(m) = \sum_{\substack{d \leq u \\ d|m}} \mu(d). \end{aligned}$$

To estimate  $S_3, S_4$  and  $S_5$ , we will consider two cases:

$$(a) |f'(X_3)| \leq \log^{-10} X \quad \text{and} \quad (b) |f'(X_3)| \geq \log^{-10} X.$$

Case (a). Assume first that  $|f'(X_3)| \leq \log^{-10} X$ . We take  $u = v = \log^3 X$ .

---

<sup>3</sup>Throughout this section we are tacitly assuming that  $n$  is a continuous variable.



Let us estimate the sum  $S_3$ . For  $X_3 \leq dm \leq X_4$ , we have

$$\left| \frac{\partial}{\partial m}(qf(dm)) \right| = |qd f'(dm)| \leq \frac{qd}{\log^{10} X} \leq \frac{1}{2} \quad \text{and} \quad \left| \frac{\partial}{\partial m}(qf(dm)) \right| = qd |f'(dm)| \geq qd \frac{\log^{15} X}{X}.$$

Using Lemma 2.5 first and Lemma 2.9 after that, we obtain

$$|S_3| \leq \sum_{d \leq u} \left| \sum_{\frac{X_3}{d} \leq m \leq \frac{X_4}{d}} \log m e(qf(dm)) \right| \quad (3.8)$$

$$\ll \log X \sum_{d \leq u} \frac{X}{qd \log^{15} X} \leq \frac{X}{\log^3 X}. \quad (3.9)$$

Let us estimate the sum  $S_4$ . For  $X_3 \leq mr \leq X_4$ , we have

$$\left| \frac{\partial}{\partial r}(qf(mr)) \right| = qm |f'(mr)| \ll \frac{qm}{\log^{10} X} \leq \frac{1}{2} \quad \text{and} \quad \left| \frac{\partial}{\partial r}(qf(mr)) \right| \geq qm \frac{\log^{15} X}{X}.$$

Using Lemma 2.9 to evaluate the sum over  $r$ , we get

$$|S_4| \leq \sum_{m \leq uv} |a(m)| \left| \sum_{\frac{X_3}{m} < r \leq \frac{X_4}{m}} e(qf(mr)) \right| \quad (3.10)$$

$$\ll \sum_{m \leq uv} |a(m)| \frac{X}{qm \log^{15} X} \ll \frac{X}{\log^{15} X} \sum_{d \leq u} \frac{|\mu(d)|}{d} \sum_{n \leq v} \frac{\Lambda(n)}{n} \ll \frac{X}{\log^3 X}. \quad (3.11)$$

To evaluate the sum  $S_5$ , we divide the interval  $(u, \frac{X_4}{u}]$  into  $\ll \log X$  subintervals of the form  $(M, M_1] \subset (M, 2M]$  and evaluate

$$S_6 := \sum_{M < m \leq M_1} \sum_{\substack{\frac{X_3}{m} \leq n \leq \frac{X_4}{m} \\ v > n}} b(m),$$

Note that

$$|S_5| \ll \log X \cdot \max |S_6|. \quad (3.12)$$

To estimate  $S_6$ , we will consider two cases  $M \geq \sqrt{X_4}$  and  $M < \sqrt{X_4}$ .

If  $M \geq \sqrt{X_4}$ , we use the Cauchy-Schwartz inequality to obtain

$$\begin{aligned} |S_6|^2 &= \left| \sum_{M < m \leq M_1} b(m) \sum_{\substack{\frac{X_3}{m} < n \leq \frac{X_4}{m} \\ n > v}} \Lambda(n) e(qf(mn)) \right|^2 \\ &\leq \left( \sum_{M < m \leq M_1} |b(m)|^2 \right) \left( \sum_{M < m \leq M_1} \left| \sum_{\substack{\frac{X_3}{m} < n \leq \frac{X_4}{m} \\ n > v}} \Lambda(n) e(qf(mn)) \right|^2 \right) \quad (3.13) \end{aligned}$$

With the help of Remark 2.2, the first term in (3.13) can be estimated as follows:

$$\sum_{M < m \leq M_1} |b(m)|^2 \ll M \log^3 u.$$

To estimate the second term in (3.13), we will apply Lemma 2.6 with the largest integer  $H \leq \frac{X_3}{M_1}$  such that

$$Hq \max_{X_3 \leq t \leq X_4} |f'(t) + tf''(t)| \leq \frac{1}{2}. \quad (3.14)$$

Then,

$$|S_6|^2 \ll M \log^3 u |A + B|, \text{ where} \quad (3.15)$$

$$A = \sum_{M < m \leq M_1} \frac{X_4}{MH} \left( \sum_{h=1}^H \sum_{\substack{\frac{X_3}{m} \leq n, n+h \leq \frac{X_4}{m} \\ n, n+h > v}} \Lambda(n) \Lambda(n+h) e(qf(mn+mh) - qf(mn)) \right)$$

$$B = \sum_{M < m \leq M_1} \frac{X_4}{MH} \sum_{\substack{\frac{X_3}{m} \leq n \leq \frac{X_4}{m} \\ n > v}} \Lambda^2(n).$$

Note that for some  $t_0 \in [t, t_1] = [mn, mn+mh]$

$$\begin{aligned} \left| \frac{\partial}{\partial m} (qf(mn+mh) - qf(mn)) \right| &= |q[(n+h)f'(mn+mh) - nf'(mn)]| \\ &= \frac{q}{m} |(t_1 f'(t_1) - t f'(t))| = \frac{q(t_1 - t)}{m} \left| \frac{d}{dt} (t f'(t)) \right|_{t=t_0} = qh |f'(t_0) + t_0 f''(t_0)|. \end{aligned}$$

For any pair  $(h, n)$ , where  $n, n+h \in (\frac{X_3}{M_1}, \frac{X_4}{M}]$ , we will use Lemma 2.9 with  $\lambda = qh \min_{X_3 \leq t \leq X_4} |f'(t) + tf''(t)|$  to estimate

$$\left| \sum_{M < m \leq M_1} e(qf(mn+mh) - qf(mn)) \right| \ll \frac{1}{\lambda}.$$

To estimate (3.15), we consider two cases  $H = \frac{X_3}{M_1}$  and  $H < \frac{X_3}{M_1}$  for (3.14). If  $H = \frac{X_3}{M_1}$ , use  $f'(n) > \frac{\log^{15} X}{X}$  for  $n \geq X_0$  and (2.4) from Proposition 2.2 to get

$$\lambda = qh |f'(t_0)| \left| 1 + \frac{t_0 f''(t_0)}{f'(t_0)} \right| \gg qh \frac{\log^{15} X}{X} \frac{1}{\log^2 X}. \quad (3.16)$$

Hence

$$\begin{aligned} |S_6|^2 &\ll M \log^3 u \sum_{h=1}^H \sum_{\substack{\frac{X_3}{M_1} < n, n+h \leq \frac{X_4}{M}}} \Lambda(n) \Lambda(n+h) \frac{X}{qh \log^{13} X} + M \log^3 u \cdot \frac{X_3}{M} \log \frac{X_3}{M} \\ &\ll \frac{X X_3 \log^3 u}{\log^{11} X} + X \log^3 u \log X \ll \frac{X^2}{\log^6 X}. \end{aligned} \quad (3.17)$$

For  $H < \frac{X_3}{M_1}$ , note that

$$|f'(X_4)| \gg \frac{|f'(X_3)|}{\log X} \gg \frac{1}{\log X} \max_{X_3 \leq t \leq X_4} |t f''(t) + f'(t)|.$$

From (2.4), (2.6) and the fact that  $(H+1)q \max_{X_3 \leq t \leq X_4} |tf''(t) + f'(t)| \geq \frac{1}{2}$ , we get

$$\begin{aligned} \lambda &= qh \min_{X_3 \leq t \leq X_4} |tf''(t) + f'(t)| \gg qh \min_{X_3 \leq t \leq X_4} \frac{|f'(t)|}{\log^2 t} \gg \frac{qh}{\log^2 X} |f'(X_4)| \\ &\gg \frac{qh}{\log^3 X} \max_{X_3 \leq t \leq X_4} |tf''(t) + f'(t)| \gg \frac{h}{2H \log^3 X}. \end{aligned} \quad (3.18)$$

Using Lemma 2.9, we can estimate  $A$ :

$$\begin{aligned} A &\ll \frac{X_4}{MH} \sum_{h=1}^H \sum_{\substack{X_3/M_1 < n, n+h \leq X_4/M}} \Lambda(n)\Lambda(n+h) \frac{2H \log^3 X}{h} \\ &\ll \frac{X_4}{MH} \left( \frac{X_3}{M_1} \log \frac{X_3}{M_1} \right) (2H \log^3 X) \log H \ll \frac{X^2 \log^5 X}{M^2}. \end{aligned}$$

Now let us estimate  $B$ . From (2.4),  $\lambda = qh \min_{X_3 \leq t \leq X_4} |f'(t) + tf''(t)| \leq qh |f'(X_3)|$ .

So from (3.18),

$$\frac{1}{H} \ll \frac{2\lambda \log^3 X}{h} \ll \frac{2 \log^3 X}{h} qh |f'(X_3)| \ll \frac{2q \log^3 X}{\log^{10} X} \leq \frac{2}{\log^6 X}. \quad (3.19)$$

Then,

$$B \ll \sum_{M < m \leq M_1} \frac{X_4}{MH} \left( \frac{X_3}{m} \log \frac{X_3}{m} \right) \ll \frac{X^2 \log X}{MH} \ll \frac{X^2}{M \log^5 X}.$$

Therefore, we get

$$\begin{aligned} |S_6|^2 &\ll \frac{X^2 \log^3 u \cdot \log^5 X}{M} + \frac{X^2 \log^3 u}{\log^5 X} \\ &\ll X^{\frac{3}{2}} \log^6 X + \frac{X^2}{\log^{9/2} X} \ll \frac{X^2}{\log^{9/2} X}. \end{aligned} \quad (3.20)$$

From (3.17) and (3.20), summing over all  $M \geq \sqrt{X_4}$ , we obtain

$$|S_5| \ll \log X \cdot \frac{X}{\log^{9/4} X} = \frac{X}{\log^{5/4} X}. \quad (3.21)$$

If  $M < \sqrt{X_4}$ , we interchange the order of summation in (3.13) and again use the Cauchy-Schwartz inequality:

$$\begin{aligned} |S_6|^2 &\leq \left| \sum_{\substack{N \leq n \leq N_1 \\ n > v}} \Lambda(n) \sum_{\substack{X_3/n \leq m \leq X_4/n}} b(m) e(qf(mn)) \right|^2 \\ &\leq \left( \sum_{N \leq n \leq N_1} \Lambda^2(n) \right) \left( \sum_{N \leq n \leq N_1} \left| \sum_{\substack{X_3/n \leq m \leq X_4/n}} b(m) e(qf(mn)) \right|^2 \right), \end{aligned}$$

where  $N = \frac{X_3}{M_1}$  and  $N_1 = \frac{X_4}{M}$ . The rest of the proof is similar to proof for the case  $M \geq \sqrt{X_4}$ .

Let  $H$  be the largest integer  $\leq M/2$  such that  $qh \max_{X_3 \leq t \leq X_4} |f'(t) + tf''(t)| \leq 1/2$ . Then,

$$|S_6|^2 \ll N \log N |C + D|, \text{ where} \quad (3.22)$$

$$C = \sum_{N \leq n \leq N_1} \frac{X_4}{NH} \left( \sum_{h=1}^H \sum_{\frac{X_3}{n} \leq m, m+h \leq \frac{X_4}{n}} b(m)b(m+h)e(qf(mn+mh) - qf(mn)) \right)$$

$$D = \sum_{N \leq n \leq N_1} \frac{X_4}{NH} \sum_{\frac{X_3}{n} \leq m \leq \frac{X_4}{n}} |b(m)|^2.$$

For any pair  $(m, h)$ , where  $m, m+h \in (\frac{X_3}{N_1}, \frac{X_4}{N}]$ , using Lemma 2.9 with  $\lambda = qh \min_{X_3 \leq t \leq X_4} |f'(t) + tf''(t)| \leq 1/2$ , we get

$$\left| \sum_{N \leq n \leq N_1} e(qf(mn+mh) - qf(mn)) \right| \ll \frac{1}{\lambda}.$$

From Remark 2.2,

$$\sum_{M \leq m \leq M_1} |b(m)|^2 \ll M \log^3 u, \quad \sum_{M \leq m, m+h \leq M_1} |b(m)| |b(m+h)| \ll M \log^3 u.$$

If  $H = \frac{X_3}{N_1} (= M/2)$ , then, as above (see (3.16)), we get  $\lambda \geq \frac{qh \log^{13} X}{X}$  and so

$$|S_6|^2 \ll N \log N \left( \sum_{h=1}^H \sum_{M \leq m, m+h \leq M_1} |b(m)| |b(m+h)| \frac{X}{qh \log^{13} X} + \sum_{M \leq m \leq M_1} |b(m)|^2 \right)$$

$$\ll \frac{X_4 X \log^3 u}{\log^{11} X} + X_4 \log \frac{X_4}{M} \log^3 u \ll \frac{X^2}{\log^6 X}. \quad (3.23)$$

If  $H < M/2$ , then, as in (3.18) and (3.19), we get  $\lambda \geq \frac{h}{2H \log^3 X}$  and  $\frac{1}{H} \ll \frac{1}{\log^6 X}$ . Thus we obtain

$$C \ll \frac{X_4}{NH} (M \log^3 u) (2H \log^3 X) \log H \ll \frac{X^2}{N^2} \log^5 X$$

and

$$D \ll \frac{X_4}{NH} \sum_{N \leq n \leq N_1} \frac{X_3}{n} \log^3 u \ll \frac{X_4}{NH} X_3 \log^3 u \ll \frac{X^2 \log^3 u}{N \log^6 X},$$

so

$$|S_6|^2 \leq \frac{X^2 \log^6 X}{N} + \frac{X^2 \log^3 u}{\log^5 X} \ll \frac{X^2}{\log^{9/2} X}. \quad (3.24)$$

From (3.23) and (3.24), summing  $|S_6|$  over all  $M < \sqrt{X_4}$  gives again the same estimate for  $S_5$  as in (3.21). Therefore,

$$|S_1| \ll \frac{\log Q}{\log X} \left( \frac{X}{\log^3 X} + \frac{X}{\log^3 X} + \frac{X}{\log^{5/4} X} \right) + \sqrt{X} \ll \frac{\pi(X)}{Q}.$$

Case (b). Now we assume that  $|f'(X_3)| \geq \log^{-10} X$ . Take  $u = v = X^{1/10}$  and argue as above but use Lemma 2.10 with  $k = 1$  instead of Lemma 2.9. The formulas

(2.1) and (2.4) imply that for  $x \in (X_3, X_4] \subset (X_3, 2X_3]$

$$|f''(x)| \ll \frac{|f'(x)|}{x} \ll \frac{1}{X_3}, \quad |f''(x)| \gg \frac{|f'(x)|}{x \log^2 x} \gg \frac{1}{X_3 \log^{13} X_3}. \quad (3.25)$$

To evaluate  $S_3$ , let  $\lambda_2 = \left| \frac{d^2}{dm^2}(qf(dm)) \right| = qd^2|f''(dm)|$ . Then

$$\lambda := \frac{qd^2}{X_3 \log^{13} X_3} \ll \lambda_2 \ll \frac{qd^2}{X_3} =: \alpha\lambda.$$

Apply (2.15) in Lemma 2.10 with  $k = 1$  to evaluate the sum in (3.8):

$$\begin{aligned} |S_3| &\ll \sum_{d \leq u} \log X \left| \sum_{\frac{X_3}{d} \leq m \leq \frac{X_4}{d}} e(qf(dm)) \right| \\ &\ll \log X \sum_{d \leq u} \frac{X_3}{d} \left( \sqrt{\frac{qd^2}{X_3}} + \sqrt{\frac{d \log^{14} X_3}{X_3}} \right) \ll X^{4/5}. \end{aligned}$$

Similarly, to evaluate  $S_4$  let  $\mu_2 = \left| \frac{d^2}{dr^2} qf(mr) \right| = qm^2|f''(mr)|$ . Then

$$\lambda := \frac{qm^2}{X_3 \log^{13} X_3} \ll \mu_2 \ll \frac{qm^2}{X_3} =: \alpha\lambda.$$

Apply Lemma 2.10 with  $k = 1$  to evaluate the sum in (3.10):

$$\begin{aligned} |S_4| &\ll \sum_{m \leq uv} |a(m)| \left| \sum_{\frac{X_3}{m} \leq r \leq \frac{X_4}{m}} e(qf(mr)) \right| \\ &\ll \sum_{m \leq uv} |a(m)| \frac{X_3}{m} \left( \sqrt{\frac{qm^2}{X_3}} + \sqrt{\frac{m \log^{14} X_3}{X_3}} \right) \ll X^{4/5}. \end{aligned}$$

To evaluate  $S_6$  for  $M \geq \sqrt{X_4}$ , we evaluate the sum over  $m$  in (3.15) with  $H = \min\{\frac{X_3}{M_1}, M^{1/3}\}$ . Denoting  $g(m) = qf(mn+mh) - qf(mn)$  with  $(t_1 = mn+mh, t = mn, t_0 \in [t, t_1])$ , as above, we get

$$\begin{aligned} g''(m) &= q[(n+h)^2 f''(mn+mh) - n^2 f''(mn)] = \frac{q}{m^2} [t_1^2 f''(t_1) - t^2 f''(t)] \\ &= \frac{q}{m^2} \cdot mh \cdot \frac{d}{dt} (t^2 f''(t)) \Big|_{t=t_0} = \frac{qh}{m} (2t_0 f''(t_0) + t_0^2 f'''(t_0)). \end{aligned}$$

Using (2.5) from Proposition 2.2 we get

$$|g''(m)| = \frac{qh}{m} |t_0 f''(t_0)| \left| 2 + \frac{t_0 f'''(t_0)}{f''(t_0)} \right| \gg \frac{qh}{M} \frac{|f'(t_0)|}{\log^2 X} \frac{1}{\log^2 X} \gg \frac{qh}{M \log^{13} X} =: \lambda$$

and

$$|g''(m)| \ll \frac{qh}{M} =: \alpha\lambda.$$

Apply Lemma 2.10 with  $k = 1$  to the sum in (3.15):

$$\begin{aligned} A &\ll \frac{X_4}{MH} \sum_{h=1}^H \sum_{\frac{X_3}{M_1} \leq n, n+h \leq \frac{X_4}{M}} \Lambda(n)\Lambda(n+h) \left( M\sqrt{\frac{qh}{M}} + M\sqrt{\frac{\log^{15} X \log M}{M}} \right) \\ &\ll \frac{X^2}{M} \log X \left( \sqrt{\frac{qH}{M}} + \sqrt{\frac{\log^{15} X \log M}{M}} \right) \ll \frac{X^2 \log^9 X}{M^{4/3}}. \end{aligned}$$

Also,  $B \ll \frac{X_3^2}{MH} \log X$ . If  $H = \frac{X_3}{M_1} \leq M^{1/3}$ , then  $1/H = \frac{M_1}{X_3} \leq \frac{1}{X_3} \frac{X_4}{u} \ll X^{-1/10}$  and if  $H = M^{1/3}$ , then  $1/H \ll X^{-1/6}$ . So  $B \ll \frac{1}{M} X^{19/10}$ . Then

$$|S_6|^2 \ll M \log^3 X (A+B) \ll \frac{X^2 \log^{12} X}{M^{1/3}} + X^{19/10} \log^3 X \ll \frac{X^2}{M^{1/6}} + X^{19/10} \leq X^{23/12}.$$

If  $M < \sqrt{X_4}$  we interchange the order of summation in (3.13) and evaluate the sum over  $n$  in (3.22) using Lemma 2.10 with  $k = 1$ , in exactly the same way as above, we obtain  $|S_6|^2 \ll X_3^2/N^{1/6} \leq X^{23/12}$ .

Thus, we have  $S_5 \ll X^{23/24} \log X$ . Combining this with  $S_3 \ll X_4^{4/5}$  and  $S_4 \ll X_4^{4/5}$ , we get

$$S_2 \ll X^{4/5} + X^{23/24} \log X$$

and we obtain

$$S_1 \ll \frac{\log Q}{\log X} (X^{4/5} + X^{23/24} \log X) + \sqrt{X} \leq \frac{\pi(X)}{Q}.$$

This proves Theorem 3.2.  $\square$

**Theorem 3.3.** *Let  $f(x) \in \mathbf{U}$  be of the type  $x^{l+}$  for some  $l \geq 1$ . Then  $(f(p))_{p \in \mathcal{P}}$  is u.d. mod 1.*

*Proof.* By Lemma 2.2, it is enough to show that  $Q(X) = \log X$  satisfies

$$S := \sum_{X_0 < p \leq X} e(qf(p)) \ll \frac{\pi(X)}{Q},$$

where  $X_0 = \frac{X}{Q}$ . Note that, by Lemma 2.5,

$$S = \sum_{X_0 < n \leq X} \frac{\Lambda(n)}{\log n} e(qf(n)) + O(\sqrt{X}) \ll \frac{1}{\log X} \max_{X_2 \leq X} \left| \sum_{X_0 < n \leq X_2} \Lambda(n) e(qf(n)) \right| + \sqrt{X}.$$

Now let us estimate  $\sum_{X_0 < n \leq X_2} \Lambda(n) e(qf(n))$ . Since  $X_0 = \frac{X}{Q}$  and  $X_2 \leq X$ , we can divide the interval  $(X_0, X_2]$  into  $\ll \log Q$  subintervals of the form  $(X_3, X_4] \subset (X_3, 2X_3]$  and evaluate the corresponding sum

$$S_2 := \sum_{X_3 < n \leq X_4} \Lambda(n) e(qf(n)).$$

By Lemma 2.7 with  $u = v = X^\epsilon$  for some small  $\epsilon > 0$ , we have  $|S_2| \leq |S_3| + |S_4| + |S_5|$ , where  $S_3, S_4$  and  $S_5$  correspond to  $T_1, T_2$  and  $T_3$  in Lemma 2.7:

$$S_3 = \sum_{d \leq u} \sum_{\frac{X_3}{d} \leq m \leq \frac{X_4}{d}} \mu(d)(\log m)e(qf(dm)), \quad (3.26)$$

$$S_4 = \sum_{m \leq uv} \sum_{\frac{X_3}{m} \leq r \leq \frac{X_4}{m}} a(m)e(qf(mr)), \quad a(m) = \sum_{d \leq u} \sum_{\substack{n \leq v \\ dn=m}} \mu(d)\Lambda(n) \quad (3.27)$$

$$S_5 = \sum_{m > u} \sum_{\substack{\frac{X_3}{m} \leq n \leq \frac{X_4}{m} \\ v > n}} b(m)\Lambda(n)e(qf(mn)), \quad b(m) = \sum_{\substack{d \leq u \\ d|m}} \mu(d). \quad (3.28)$$

We will show that  $|S_i| \ll X^{1-\epsilon_0}$  for some  $\epsilon_0 > 0$ , which implies that

$$|S| \ll \frac{\log Q}{\log X} X^{1-\epsilon_0} \leq \frac{\pi(X)}{Q}.$$

Let us estimate  $S_3$  first. Let  $\lambda_j := \left| \frac{\partial^{l+j}}{\partial m^{l+j}} qf(dm) \right| = qd^{l+j} |f^{(l+j)}(dm)|$ . From Proposition 2.3, for  $dm \in (X_3, X_4]$ ,

$$qd^{l+j} X_3^{\beta-l-j-\epsilon} \ll \lambda_j \ll qd^{l+j} X_3^{\beta-l-j+\epsilon}, \quad (3.29)$$

where  $\beta = \lim_{x \rightarrow \infty} \frac{\log f(x)}{\log x} \in [l, l+1]$ . Choose  $j$  such that  $qd^{l+j} X_3^{\beta-j-l+\epsilon} \ll X_3^{-\epsilon}$ .

We will use Lemma 2.10 with  $k = l+j-1$ ,  $\lambda = qd^{k+1} X_3^{\beta-k-1-\epsilon}$  and  $\alpha\lambda = qd^{k+1} X_3^{\beta-k-1+\epsilon}$ . Note that

$$\alpha\lambda \ll X_3^{-\epsilon}, \quad \lambda \left( \frac{X_3}{d} \right)^{k+1} \gg X_3^\epsilon \quad \text{and} \quad \frac{\alpha \log^k X_3/d}{X_3/d} \ll X_3^{-\epsilon}.$$

Using Lemmas 2.5 and 2.10 we get:

$$|S_3| \ll \sum_{d \leq u} \log X_3 \frac{X_3}{d} X_3^{-\frac{\epsilon}{2K}} \ll X^{1-\frac{\epsilon}{4K}}.$$

To evaluate  $S_4$ , we will use an argument similar to that utilized in evaluating  $S_3$ . Denote

$$\mu_j := \left| \frac{\partial^{l+j}}{\partial r^{l+j}} (qf(mr)) \right| = qm^{l+j} |f^{(l+j)}(mr)|.$$

Then, similar to (3.29),

$$qm^{l+j} X_3^{\beta-l-j-\epsilon} \ll \mu_j \ll qm^{l+j} X_3^{\beta-l-j+\epsilon}, \quad (3.30)$$

so by choosing  $j$  such that  $qm^{l+j} X_3^{\beta-j-l+\epsilon} \ll X_3^{-\epsilon}$  and  $k = l+j-1$ ,

$$|S_4| \ll \sum_{m \leq uv} |a(m)| \frac{X_3}{m} X_3^{-\frac{\epsilon}{2K}} \ll \sum_{m \leq uv} \frac{|a(m)|}{m} X_3^{1-\frac{\epsilon}{2K}} \ll X^{1-\frac{\epsilon}{4K}}.$$

In order to evaluate  $S_5$ , we will estimate  $S_6$  as in the proof of Theorem 3.2. So we need to evaluate the sum in (3.15) for  $M \geq \sqrt{X_4}$  and the sum in (3.22) for  $M < \sqrt{X_4}$ .

For the sum in (3.15), take  $H = \frac{1}{2}X_3^\epsilon$ . Note that  $\frac{X_3}{M_1} \geq \frac{X_4}{2M_1} \geq \frac{1}{2}u \geq H$ . Then

$$B \ll \sum_{M < m \leq M_1} \frac{X_3}{MH} \frac{X_3}{m} \log X_3 \ll \frac{X_3^2 \log X_3}{MH} \ll \frac{X_3^{2-\epsilon} \log X_3}{M} \ll \frac{X^{2-\epsilon} \log X}{M}.$$

To evaluate  $A$ , we need to estimate

$$\left| \sum_{M < m \leq M_1} e(qf(mn + mh) - qf(mn)) \right|, \quad (3.31)$$

where  $n, n + h \in (\frac{X_3}{M_1}, \frac{X_4}{M}]$ . Using Proposition 2.3, we get

$$\begin{aligned} \lambda_{1,j} &:= \left| \frac{\partial^{l+j}}{\partial m^{l+j}} (qf(mn + mh) - qf(mn)) \right| \\ &= \frac{qh}{m^{l+j-1}} t_0^{l+j-1} \left| (l+j)f^{(l+j)}(t_0) + t_0 f^{(l+j+1)}(t_0) \right| \\ &= \frac{qh}{m^{l+j-1}} t_0^{l+j-1} \left| f^{(l+j)}(t_0) \right| \left| l+j + \frac{t_0 f^{(l+j+1)}(t_0)}{f^{(l+j)}(t_0)} \right| \end{aligned} \quad (3.32)$$

for some  $t_0 \in [mn, mn + mh]$ , so

$$\frac{qh}{M^{l+j-1}} X_3^{\beta-1-\epsilon} \ll \lambda_{1,j} \ll \frac{qh}{M^{l+j-1}} X_3^{\beta-1+\epsilon}. \quad (3.33)$$

Fix an integer  $j$  such that  $\frac{qh}{M^{l+j-1}} X_3^{\beta-1+\epsilon} \ll X_3^{-\epsilon}$ . We will evaluate (3.31) via Lemma 2.10 with  $k = l + j - 1$  and  $\lambda = \frac{qh}{M^{l+j-1}} X_3^{\beta-1-\epsilon}$  and  $\alpha\lambda = \frac{qh}{M^{l+j-1}} X_3^{\beta-1+\epsilon}$ . Note that  $\alpha\lambda \leq X_3^{-\epsilon}$ ,  $\lambda M^{k+1} \gg X_3^\epsilon$  and  $\frac{\alpha \log^k M}{M} \ll X_3^{-\epsilon}$ . Then,

$$A \ll \frac{X_4}{MH} \sum_{h=1}^H \sum_{\frac{X_3}{M_1} \leq n, n+h \leq \frac{X_4}{M}} \Lambda(n) \Lambda(n+h) M X_3^{-\epsilon/2K} \ll \frac{X^{2-\epsilon/2K} \log X}{M}.$$

Thus  $|S_6|^2 \ll X^{2-\epsilon/2K} \log^4 X$ .

For the sum in (3.22) for  $M < \sqrt{X_4}$ , let  $N = \frac{X_3}{M_1}$  and  $N_1 = \frac{X_4}{M}$ . Take  $H = \frac{1}{2} X_3^\epsilon$  so that

$$D \ll \sum_{N < n \leq N_1} \frac{X_4}{NH} \frac{X_3}{n} \log^3 u \ll \frac{X_3^{2-\epsilon} \log^3 X_3}{N} \ll \frac{X^{2-\epsilon} \log^3 X}{N}.$$

Denote

$$\begin{aligned} \mu_{1,j} &:= \left| \frac{\partial^{l+j}}{\partial n^{l+j}} (qf(mn + mh) - qf(mn)) \right| = qm^{l+j} \cdot mh \left| f^{(l+j+1)}(t_0) \right| \\ &= \frac{qh}{n^{l+j+1}} t^{l+j+1} \left| f^{(l+j+1)}(t_0) \right|, \end{aligned} \quad (3.34)$$

for some  $t_0 \in [mn, mn + mh]$ . Thus,

$$\frac{qh}{N^{l+j+1}} X_3^{\beta-\epsilon} \ll \mu_{1,j} \ll \frac{qh}{N^{l+j+1}} X_3^{\beta+\epsilon}. \quad (3.35)$$

Fix  $j$  such that  $\frac{qh}{N^{l+j+1}} X_3^{\beta+\epsilon} \ll X_3^{-\epsilon}$ . Applying Lemma 2.10 with  $k = l + j - 1$  as above to evaluate  $A$ , we also get

$$C \ll \frac{X^{2-\epsilon/2K} \log X}{N}.$$

Thus  $|S_6|^2 \ll X^{2-\epsilon/2K} \log^4 X$ , so  $|S_5| \ll X^{1-\epsilon/8K}$ .  $\square$

**Theorem 3.4.** *Let  $f_1(x) \in \mathbf{U}$  is a subpolynomial function such that*

- $f_1$  is of the type  $x^{0+}$  and  $\lim_{x \rightarrow \infty} \left| \frac{f_1(x)}{\log x} \right| = \infty$  or



- $f_1$  is of the type  $x^{l+}$  for some  $l \geq 1$ .

Let  $f(x) = f_1(x) + f_2(x)$ , where  $f_2(x) \in \mathbb{R}[x]$  is a polynomial. Then  $(f(p))_{p \in \mathcal{P}}$  is u.d. mod 1.

*Proof.* We assume that  $f_2(0) = 0$  and write  $f_2(x) = \alpha_1 x + \cdots + \alpha_t x^t \in \mathbb{R}[x]$ . If  $t \leq l$ , then  $f(x) \in \mathbf{U}$  is of the type  $x^{l+}$ , so the result follows from Theorems 3.2 and 3.3. So we also assume that  $t \geq l + 1$ .

Let us consider first  $l \geq 1$ . Note that for  $j \geq t - l + 1$ ,  $f_2^{(j+l)}(x) = 0$ . Moreover by choosing  $j$  appropriately as in the proof of Theorem 3.3 with this additional condition, we get the same estimates for  $\lambda_j, \mu_j, \lambda_{1,j}, \mu_{1,j}$  as in (3.29), (3.30), (3.33) and (3.35). (Here  $\beta = \lim_{x \rightarrow \infty} \frac{\log f_1(x)}{\log x}$ .) Hence, we can evaluate  $S_3, S_4$  and  $S_5$  as in the proof of Theorem 3.3, which concludes the proof for  $l \geq 1$ .

Now consider  $l = 0$ . We will consider two cases:

Case I. Suppose that  $|f_1'(x)| \geq \frac{\log^c x}{x}$  eventually, where  $c \geq 9 \cdot 2^t + 3t + 4$ . As in the proof of Theorem 3.3, with  $u = v = X^\epsilon$  for some small  $\epsilon > 0$ , we will show that  $S_3, S_4$  and  $S_5$  in (3.26) - (3.28) are  $O\left(\frac{X}{\log^2 X}\right)$ .

Since  $\frac{1}{\log^2 x} \ll \left| \frac{x f_1^{(j+1)}(x)}{f_1^{(j)}(x)} \right| \ll 1$ , from (2.5), we have

$$\frac{|f_1'(x)|}{x^j \log^{2j} x} \ll |f_1^{(j+1)}(x)| \ll \frac{|f_1'(x)|}{x^j}.$$

From (2.6), for  $x \in (X_3, X_4] \subset (X_3, 2X_3]$ ,

$$\frac{|f_1'(X_3)|}{X_3^j \log^{2j+1} X_3} \ll |f_1^{(j+1)}(x)| \ll \frac{|f_1'(X_3)|}{X_3^j}.$$

To evaluate  $S_3$ , we need to estimate  $\lambda_j := \left| \frac{\partial^j}{\partial m^j} (qf(dm)) \right| = qd^j |f_1^{(j)}(dm)|$  for  $j = t + 1 \geq 2$ :

$$\frac{qd^j |f_1'(X_3)|}{X_3^{j-1} \log^{2j-1} X_3} \ll \lambda_j \ll \frac{qd^j |f_1'(X_3)|}{X_3^{j-1}}.$$

We will apply Lemma 2.10 with

$$k = j - 1 (= t), \quad \lambda = \frac{qd^{k+1} |f_1'(X_3)|}{X_3^k \log^{2k+1} X_3} \quad \text{and} \quad \alpha\lambda = \frac{qd^{k+1} |f_1'(X_3)|}{X_3^k}.$$

Note that  $\alpha\lambda \ll X_3^{-\epsilon}$ ,  $\lambda \left(\frac{X_3}{d}\right)^{k+1} \gg \log^{c-(2k+1)} X_3$  and  $\frac{\alpha \log^k X_3/d}{X_3/d} \ll X_3^{-\epsilon}$ . Thus,

$$\begin{aligned} |S_3| &\ll \sum_{d \leq X_3^\epsilon} \log X \left[ \frac{X_3}{d} \log^{(-c+3k+1)/K} X_3 \right] \\ &\ll X_3 \log^2 X \log^{(-c+3k+2)/K} X_3 \ll \frac{X}{\log^2 X}. \end{aligned}$$

We can evaluate  $S_4$  similarly: for  $j = t + 1$ , denote  $\mu_j := \left| \frac{\partial^j}{\partial r^j} (qf(mr)) \right| = qm^j |f_1^{(j)}(mr)|$ . Then

$$\frac{qm^j |f_1'(X_3)|}{X_3^{j-1} \log^{2j-1} X_3} \ll \mu_j \ll \frac{qm^j |f_1'(X_3)|}{X_3^{j-1}}.$$

Applying Lemma 2.10 with

$$k = j - 1 (= t), \quad \lambda = \frac{qm^{k+1}|f'_1(X_3)|}{X_3^k \log^{2k+1} X_3} \quad \text{and} \quad \alpha\lambda = \frac{qm^{k+1}|f'_1(X_3)|}{X_3^k},$$

we have

$$|S_4| \ll \sum_{m \leq uv} |a(m)| \frac{X}{m} \log^{(-c+3k+1)/K} X_3 \ll X \log^3 X \log^{(-c+3k+1)/K} X_3 \ll \frac{X}{\log^2 X}.$$

To evaluate the sum  $S_5$ , we need to estimate  $S_6$  as in the proof of Theorem 3.2: to evaluate the sum in (3.15) for  $M \geq \sqrt{X_4}$  and the sum in (3.22) for  $M < \sqrt{X_4}$ . For  $M \geq \sqrt{X_4}$ , take  $H = \frac{X_3}{M_1}$ . Then from (3.15),  $B \ll X \log X$ . To evaluate  $A$ , for  $j = t + 1$ , denote

$$\lambda_{1,j} := \left| \frac{\partial^j}{\partial m^j} (qf(mn + mh) - qf(mn)) \right|.$$

Then, from (3.32) in the proof of Theorem 3.3,

$$\frac{qh}{M^{j-1}} \frac{|f'_1(X_3)|}{\log^{2j+1} X_3} \ll \lambda_{1,j} \ll \frac{qh}{M^{j-1}} |f'_1(X_3)|.$$

We apply now Lemma 2.10 with

$$k = j - 1 (= t), \quad \lambda = \frac{qh}{M^k} \frac{|f'_1(X_3)|}{\log^{2k+3} X_3} \quad \text{and} \quad \alpha\lambda = \frac{qh}{M^k} |f'_1(X_3)|.$$

Note that

$$\alpha\lambda \ll X_3^{-\epsilon}, \quad \frac{\alpha \log^k M}{M} \ll X_3^{-\epsilon}$$

$$\text{and } (\lambda M^{k+1})^{-1/K} (\log M)^{k/K} \ll \left( \frac{X_3}{hM} \right)^{1/K} \log^{-9} X_3.$$

Hence

$$\begin{aligned} A &\ll \frac{X_4}{MH} \sum_{h=1}^H \left( \frac{1}{h} \right)^{1/K} \sum_{\frac{X_3}{M_1} < n, n+h \leq \frac{X_4}{M}} \left( M \left( \frac{X_3}{M} \right)^{1/K} \log^{-9} X_3 \right) \\ &\ll \frac{X_4}{MH} H^{1-1/K} \frac{X_3}{M_1} \log X_3 \left( M \left( \frac{X_3}{M} \right)^{1/K} \log^{-9} X_3 \right) \ll \frac{X^2}{M} \log^{-8} X, \end{aligned}$$

thus in this case

$$|S_6|^2 \ll M \log^3 u |A + B| \ll M \log^3 X |A + B| \ll X^2 \log^{-8} X.$$

Similarly, for  $M < \sqrt{X_4}$ , take  $H = M/2 = \frac{X_3}{N_1}$ . Then from (3.22)  $D \ll X \log^3 X$ . To evaluate  $C$ , for  $j = t + 1$ , denote

$$\mu_{1,j} := \left| \frac{\partial^j}{\partial n^j} (qf(mn + mh) - qf(mn)) \right|.$$

Then, from (3.34) in the proof of Theorem 3.3,

$$\frac{qh}{N^{j+1}} X_3 \frac{|f'_1(X_3)|}{\log^{2j+1} X_3} \ll \mu_{1,j} \ll \frac{qh}{N^{j+1}} X_3 |f'_1(X_3)|.$$

Applying again Lemma 2.10 with  $k = j - 1 (= t)$ , we get

$$\lambda = \frac{qh}{N^{k+2}} X_3 \frac{|f'_1(X_3)|}{\log^{2k+3} X_3} \text{ and } \alpha\lambda = \frac{qh}{N^{k+2}} X_3 |f'_1(X_3)| \text{ for } C.$$

Note that

$$\alpha\lambda \ll X_3^{-\epsilon}, \quad \frac{\alpha \log^k N}{N} \ll X_3^{-\epsilon}$$

$$\text{and } (\lambda N^{k+1})^{-1/K} (\log N)^{k/K} \ll \left(\frac{N}{h}\right)^{1/K} \log^{-9} X_3,$$

thus

$$C \ll \frac{X_4}{NH} H^{1-1/K} \frac{X_3}{N_1} \log^3 X_3 \left(N^{1+1/K} \log^{-9} X_3\right)$$

so in this case,

$$|S_6|^2 \ll N \log N (C + D) \ll X^2 \log^{-5} X.$$

Hence we have  $|S_5| \ll X \log^2 X$ . So we are done with Case I.

Case II. Assume that  $|f'_1(x)| \ll \frac{\log^c x}{x}$  eventually for some  $c > 0$ .

If all coefficients of  $f_2(x)$  are rational, write  $f_2(x) = \frac{1}{r} \sum_{j=1}^t a_j x^j$ , where  $a_1, \dots, a_t, r$  are integers. From Lemma 2.2, we will prove that

$$\left| \sum_{X_0 < p \leq X} e(qf_1(p) + qf_2(p)) \right| \ll \frac{\pi(X)}{Q}, \quad (3.36)$$

where  $Q = Q(X) \leq \log X$  is a positive unbounded increasing function with  $Q(X) \leq X |f'_1(X)|$ . Note that

$$\left| \sum_{X_0 < p \leq X} e(qf_1(p) + qf_2(p)) \right| \ll \sum_{\substack{a=1 \\ (a,r)=1}}^r \left| \sum_{\substack{X_0 < p \leq X \\ p \equiv a \pmod{r}}} e(qf_1(p)) \right|.$$

Write

$$S(a) := \left| \sum_{\substack{X_0 < p \leq X \\ p \equiv a \pmod{r}}} e(qf_1(p)) \right| = \sum_{X_0 < n \leq X} \Lambda_1(n, r, a) e(qf_1(n)).$$

Applying Lemma 2.4 and arguing as in the proof of the estimate for  $S_0$  in (3.4) above we obtain

$$|S(a)| \ll \frac{1}{\log X} \frac{1}{q |f'_1(X)|} + \frac{\pi(X)}{Q} \ll \frac{\pi(X)}{Q}.$$

So this takes care of the case when all coefficient of  $f_2(x)$  are rational.

Now it remains to consider the case when at least one coefficient of  $f_2(x)$  is irrational. We will show that for any given non-zero integer  $q$ ,

$$\lim_{X \rightarrow \infty} \frac{1}{\pi(X)} \left| \sum_{X_0 < p \leq X} e(qf_1(p) + qf_2(p)) \right| = 0, \quad (3.37)$$

which obviously implies

$$\lim_{X \rightarrow \infty} \frac{1}{\pi(X)} \left| \sum_{p \leq X} e(qf_1(p) + qf_2(p)) \right| = 0.$$

For any positive integer  $Y$ , write all coefficients of  $f_2(x)$  in the form

$$\begin{aligned} \alpha_j &= \frac{e_j}{q_j} + z_j \quad \text{with} \quad (e_j, q_j) = 1, \quad 0 < q_j \leq \tau_j := Y^{j/2} \\ |z_j| &\leq \frac{1}{q_j \tau_j}, \quad z_j = \delta_j Y^{-j} \quad (j = 1, 2, \dots, t). \end{aligned} \quad (3.38)$$

Denote

$$\delta_0 = \max\{|\delta_1|, \dots, |\delta_t|\} \quad \text{and} \quad \mathcal{Q} = [q_1, \dots, q_t]. \quad (3.39)$$

Assume first that there exists some  $Y$  with  $X_0 \leq Y \leq X$ , ( $X_0 = \frac{X}{\log X}$ ), such that

$$\delta_0(Y) \leq \exp((\log X)^{1/3}) \quad \text{and} \quad \mathcal{Q}(Y) \leq \exp((\log \log X)^3). \quad (3.40)$$

To simplify our notation, we use  $\delta_0, \mathcal{Q}$  instead of  $\delta_0(Y)$  and  $\mathcal{Q}(Y)$ . Also, denote by  $\sum^*$  the sum over all  $a$  from the reduced residue system modulo  $\mathcal{Q}$ . We obtain

$$S = \sum_{X_0 < p \leq X} e(qf_1(p) + qf_2(p)) = \sum^* \sum_{X_0 < n \leq X} \Lambda_1(n, \mathcal{Q}, a) e(qf_1(n) + qf_2(n)).$$

Denoting  $h(n) = \sum_{j=1}^t z_j n^j$ , we get

$$S = \sum_a^* e \left( \sum_{j=1}^k \frac{e_j}{q_j} a^j \right) \sum_{X_0 < n \leq X} \Lambda_1(n, \mathcal{Q}, a) e(qf_1(n) + qh(n)). \quad (3.41)$$

Now we will evaluate  $\sum_{n=X_0}^{X_1} \Lambda_1(n, \mathcal{Q}, a) e(qf_1(n) + qh(n))$  using the second part of Lemma 2.5 with  $b_n = \Lambda_1(n, \mathcal{Q}, a)$  and  $a_n = e(qf_1(n) + qh(n))$ . Lemma 2.3 implies that

$$\sum_{n \leq x} b_n = \frac{1}{\phi(\mathcal{Q})} Li(x) - E \chi(a) \int_2^x \frac{u^{\beta_1-1}}{\log u} du + R(x)$$

with  $|R(x)| \ll X \exp(-C(\log X)^{1/2}) = R$ . Also,  $|a_n| = 1$  and

$$\begin{aligned} \sum_{n \leq X} |a_n - a_{n+1}| &= \sum_n |e(qf_1(n) - qf_1(n+1) + qh(n) - qh(n+1)) - 1| \\ &\ll \sum_n |q|f_1(n) - f_1(n+1)| + |q| h(n) - h(n+1)| \\ &\ll q(\log X)^{c+1} + q \sum |z_j| X^j \ll q(\log X)^{c+1} + q \sum |z_j| Y^j \log^j X \\ &\ll q((\log X)^{c+1} + \delta_0 \log^t X) \ll \exp((\log X)^{\frac{1}{3} + \epsilon_1}) \end{aligned}$$

for some small  $\epsilon_1 > 0$ . Using this and Lemma 2.5 we obtain

$$\begin{aligned} \sum a_n b_n &= \frac{1}{\phi(\mathcal{Q})} \sum_n \left( Li(n) - Li(n-1) - E \chi(a) \int_{n-1}^n \frac{u^{\beta_1-1}}{\log u} du \right) e(qf_1(n) + qh(n)) \\ &\quad + O \left( X \exp \left( (\log X)^{\frac{1}{3} + \epsilon_1} - C(\log X)^{1/2} \right) \right). \end{aligned}$$

Here  $Li(n) - Li(n-1) = \frac{1}{\log n} + O\left(\frac{1}{n \log n}\right)$  and  $\int_{n-1}^n \frac{u^{\beta_1-1}}{\log u} du = \frac{n^{\beta_1-1}}{\log n} + O\left(\frac{1}{n \log n}\right)$ . So, we get

$$\begin{aligned} \sum a_n b_n &= \frac{1}{\phi(Q)} \sum_n \frac{1}{\log n} (1 - E\chi(a)n^{\beta_1-1}) e(qf_1(n) + qh(n)) + O\left(X \exp\left(-\frac{C}{2}(\log X)^{1/2}\right)\right) \\ &= \frac{1}{\phi(Q)} \sum_n \frac{1}{\log n} e(qf_1(n) + qh(n)) - \frac{E\chi(a)}{\phi(Q)} \sum_n \frac{n^{\beta_1-1}}{\log n} e(qf_1(n) + qh(n)) \quad (3.42) \\ &\quad + O\left(X \exp\left(-\frac{C}{2}(\log X)^{1/2}\right)\right). \end{aligned}$$

By using the definition of  $h(n)$  above, we can write

$$f_2(n) - h(n) = \sum_{j=1}^t \frac{e_j}{q_j} a^j \pmod{1} \quad \text{for } n = a \pmod{Q}.$$

Using summation by parts on each of the sums in (3.42) above and putting the obtained expression into (3.41), we obtain

$$\begin{aligned} S &\ll \frac{1}{\log X} \max_{X_2 \in (X_0, X]} \left| \sum_{n \in (X_0, X_2]} e(qf_1(n) + qf_2(n)) \right| + \frac{X}{\log^2 X} \\ &\ll \frac{1}{\log X} \left( \frac{X}{\log X} + \max_{X_2 \in (X_0, X]} \left| \sum_{n \leq X_2} e(qf_1(n) + qf_2(n)) \right| \right) + \frac{X}{\log^2 X}. \quad (3.43) \end{aligned}$$

By Boshernitzan's theorem,

$$\lim_{Y \rightarrow \infty} \frac{1}{Y} \left| \sum_{n \leq Y} e(qf_1(n) + qf_2(n)) \right| = 0,$$

therefore, from (3.43)

$$\lim_{X \rightarrow \infty} \frac{1}{\pi(X)} \left| \sum_{X_0 < p \leq X} e(qf_1(p) + qf_2(p)) \right| = \lim_{X \rightarrow \infty} \left| \frac{S}{\pi(X)} \right| = 0.$$

Finally, assume that for each  $Y$  with  $X_0 \leq Y \leq X$ ,  $(X_0 = \frac{X}{\log X})$ ,

$$\delta_0(Y) \geq \exp((\log X)^{1/3}) \quad \text{or} \quad Q(Y) \geq \exp((\log \log X)^3). \quad (3.44)$$

For (3.36), we use the partial summation formula with  $a_n = e(qf_1(n))$  for  $n \geq X_0$  and  $a_n = 0$  for  $n < X_0$  and  $b_n = \Lambda_1(n)e(qf_2(n))$ :

$$S = \left| \sum_p e(qf_1(p) + qf_2(p)) \right| \ll qX \max_{n \in [X_0, X]} |f_1(n+1) - f_1(n)| \max_{X_2 \leq X} \left| \sum_{p \leq X_2} e(qf_2(p)) \right|.$$

Now we use Lemma 2.12 to evaluate the last exponential sum. Since  $f_2(x) = f_2(x) + x^{12} \pmod{1}$ , we can assume that  $t \geq 12$ . We claim that

$$\Delta_1 \ll \exp(-(\log \log X)^3 / (3t)). \quad (3.45)$$

For  $Q \leq X^{\nu/5}$  and  $\delta_0 \leq X^\nu$  (recall that  $\nu = 1/t$ ), if  $\delta_0 \geq 1$ , then  $\Delta_1 = \delta_0^{-\nu/2} Q^{-\nu/2+\epsilon_0}$  satisfies (3.45) and if  $\delta_0 < 1$ , for  $m \leq X$ ,  $\Delta_1 = (m, Q)^{\nu/2} Q^{-\nu/2+\epsilon_0}$  satisfies (3.45) too. For  $Q \geq X^{\nu/5}$  or  $\delta_0 \geq X^\nu$  the claim also follows from

$X^p \gg X^{1/(20t^3)}$ . From (3.45),  $(\Delta_1)^{-2} \geq \log X$  for sufficiently large  $X$ , so we can estimate  $\left| \sum_{p \leq X_2} e(qf_2(p)) \right|$  via Lemma 2.12. Using this and

$$q \max_{n \in [X_0, X_1]} |f_1(n+1) - f_1(n)| \ll qf'_1(X_0) \ll \frac{\log^{c+2} X}{X},$$

we get

$$S \ll X(\log X)^{c+2} \exp\left(-(\log \log X)^3/(3t) + (\log \log X)^2/\log(1+\epsilon_0)\right) \ll \frac{\pi(X)}{\log X}.$$

This completes the proof.  $\square$

**Theorem 3.5.** *Let  $f_1(x)$  be a  $C^1$ -function such that for some  $C > 0$ ,  $|xf'_1(x)| \leq C$  for all  $x$  and let  $f_2(x) = \alpha_1 x + \dots + \alpha_k x^k$  be a polynomial with at least one irrational coefficient. Put  $f(x) = f_1(x) + f_2(x)$ . Then  $(f(p))_{p \in \mathcal{P}}$  is u.d. mod 1.*

*Proof.* By Lemma 2.2 it is enough to prove that there exists some  $Q := Q(X)$  for large  $X$  such that  $Q(X) \leq \log X$ ,  $Q(X) \uparrow \infty$  and for any  $q \leq Q$ ,

$$S := \left| \sum_{X_0 < p \leq X} e(qf(p)) \right| \ll \frac{\pi(X)}{Q}, \quad (3.46)$$

where  $X_0 = \frac{X}{Q}$ .

The estimation of the last sum is very similar to the estimation in Theorem 3.4. As in (3.38), for any  $\frac{X}{\log^4 X} \leq Y \leq X$ , we obtain  $\delta_0 = \delta_0(Y)$  and  $\mathcal{Q} = \mathcal{Q}(Y)$  as following:

$$\begin{aligned} \alpha_j &= \frac{e_j}{q_j} + z_j \quad \text{with} \quad (e_j, q_j) = 1, \quad 0 < q_j \leq \tau_j := Y^{j/2} \\ |z_j| &\leq \frac{1}{q_j \tau_j}, \quad z_j = \delta_j Y^{-j} \quad (j = 1, 2, \dots, k) \\ \delta_0 &= \max\{\delta_1, \dots, \delta_k\} \quad \text{and} \quad \mathcal{Q} = [q_1, \dots, q_k] \end{aligned} \quad (3.47)$$

We will treat separately the following three cases:

- (1)  $\delta_0 \mathcal{Q} \geq \exp((\log \log X)^3)$  for any  $\frac{X}{\log^4 X} \leq Y \leq X$
- (2)  $\mathcal{Q} \geq \exp((\log \log X)^3)$  for any  $\frac{X}{\log^4 X} \leq Y \leq X$
- (3)  $\delta_0 \mathcal{Q} \leq \exp((\log \log X)^3)$  and  $\mathcal{Q} \leq \exp((\log \log X)^3)$  for some  $\frac{X}{\log^4 X} \leq Y \leq X$

Let us first consider cases (1) and (2). Using Lemma 2.5 with  $a_n = e(qf(n))$  for  $n \geq X_0$  and  $a_n = 0$  for  $n < X_0$  and  $b_n = \Lambda_1(n)e(qf_2(n))$ , for any  $Z$ , we obtain

$$\left| \sum_{Z \leq n \leq 2Z} e(qf(p)) \right| \ll q \max_{X_1 \leq 2Z} \left| \sum_{n \leq X_1} \Lambda_1(n)e(qf_2(n)) \right|,$$

since  $|a_n| = 1$  and  $\sum_{Z \leq n \leq 2Z} |a_{n+1} - a_n| \ll q$ . Thus, from (3.46), dividing the interval  $(X_0, X]$  into  $\log Q$  subintervals of the form  $(Z, 2Z]$ , we obtain

$$S \ll q \log Q \max_{X_1 \leq X} \left| \sum_{n \leq X_1} \Lambda_1(n)e(qf_2(n)) \right|.$$

If  $X_1 \leq \frac{X}{\log^4 X}$ , then we have

$$q \log Q \left| \sum_{n \leq X_1} \Lambda_1(n) e(qf_2(n)) \right| \leq q \log Q \pi(X_1) \ll \frac{X}{\log^3 X}. \quad (3.48)$$

Assume now that  $X_1 \geq \frac{X}{\log^4 X}$ . If  $\mathcal{Q}(X_1) > X^{1/5k}$  or  $\delta_0(X_1) > X^{1/k}$  or if  $\mathcal{Q}(X_1) \leq X^{1/5k}$  and  $\delta_0(X_1) \leq X^{1/k}$  but (1) or (2) are satisfied then the number  $\Delta_1$  from Lemma 2.12 is  $\ll \exp((\log \log X)^3(\epsilon - 1/2k))$  for some small  $\epsilon > 0$ . Then

$$q \log X \left| \sum_{n \leq X_1} \Lambda_1(n) e(qf_2(n)) \right| \ll \frac{X}{\log^3 X}.$$

This gives  $S \ll \frac{X}{\log^3 X}$ .

We turn now our attention to the case (3). Note that  $\mathcal{Q}(Y) \leq \exp(\log \log X)^3$  for some  $Y$ . As we did in Theorem 3.4 for the case in (3.40), we get the same inequality as in (3.43):

$$S \ll \frac{1}{\log X} \max_{X_1 \in (X_0, X]} \left| \sum_{X_0 < n \leq X_1} e(qf(n)) \right| + \frac{X}{\log^2 X}. \quad (3.49)$$

We can assume that  $\alpha_k$ , the leading coefficient of  $f_2(n)$ , is irrational, (otherwise, we can divide the interval  $(X_0, X]$  in the sum (3.49) into residue classes and remove  $\alpha_k n^k$ .) We will show (3.46) with the additional restriction  $\mathcal{Q}(X) \leq \left( q_k \left( \frac{X}{\log^4 X} \right) \right)^{1/4K}$ . Note that  $\mathcal{Q}(X) \leq (q_k(Y))^{1/4K}$ . For convenience of notation, let  $q_k := q_k(Y)$ .

We denote the sum  $\left| \sum_{X_0 < n \leq X_1} e(qf(n)) \right|$  in (3.49) by  $S_1$ . Let us first consider  $k > 1$ . Using Lemma 2.6 with  $H_1 = (q_k)^{1/2K}$  we obtain

$$\left| \frac{S_1}{X_1} \right|^2 \ll \frac{1}{H_1} + \frac{1}{X_1 H_1} \sum_{h_1=1}^{H_1} \left| \sum_n e(qf(n+h_1) - qf(n)) \right|. \quad (3.50)$$

Here  $|qf_1(n+h_1) - qf_1(n)| \ll qh_1/X_0$  so if we remove  $qf_1(n+h_1) - qf_1(n)$  from the last exponential sum, we will make an error  $\ll qH_1/X_0 \ll 1/H_1$ , which allows us to replace  $f$  with  $f_2$  in (3.50).

Applying Lemma 2.6 ( $k-2$ ) more times with  $H_j = (H_1)^{J/2}$  ( $j = 1, \dots, k-1$ ), where  $J = 2^j$  and  $K = 2^k$  we obtain

$$\left| \frac{S_1}{X_1} \right|^{K/2} \ll \frac{1}{H_1^{K/4}} + \dots + \frac{1}{H_{k-1}} + \frac{1}{X_1 H_1 \dots H_{k-1}} \sum_{h_1=1}^{H_1} \dots \sum_{h_{k-1}=1}^{H_{k-1}} \left| \sum_n e(qp_1(n)) \right|, \quad (3.51)$$

where  $p_1(x) = h_1 \dots h_{k-1} \int_0^1 \dots \int_0^1 \frac{\partial^{k-1}}{\partial x^{k-1}} f_2(x + h_1 t_1 + \dots + h_{k-1} t_{k-1}) dt_1 \dots dt_{k-1}$ . Here  $\left\| \frac{dp_1(x)}{dx} \right\| = \|\alpha_k k! h_1 \dots h_{k-1}\| \gg 1/q_k$ , so if  $q_k$  is large enough that  $q \leq q_k^{1/4}$ , then  $\left\| q \frac{dp_1(x)}{dx} \right\| \gg \frac{1}{q_k}$ . Thus by Lemma 2.9 we obtain  $|\sum_n e(qp_1(n))| \ll q_k$ , which implies

$$\left| \frac{S_1}{X_1} \right|^{K/2} \ll q_k^{-1/8} + \frac{q_k}{X_1} \leq q_k^{-1/8}$$

since  $q_k \leq Q \leq \exp((\log \log X)^3)$ . Thus  $S_1 \leq X_1 q_k^{-1/4K} \ll X/Q$ , so we have  $S \ll \frac{\pi(X)}{Q}$ .

Now consider  $k = 1$ . We divide the interval  $(X_0, X_1]$  in (3.49) into  $\log Q$  subintervals of the form  $(Z, 2Z]$  and apply Lemma 2.5 with  $a_n = e(qf_1(n))$  and  $b_n = e(qf_2(n))$  to obtain

$$\left| \sum_{X_0 < n \leq X_1} e(qf(n)) \right| \ll q \log Q \max_{X_2 \leq X_1} \left| \sum_{X_0 \leq n \leq X_2} e(qf_2(n)) \right|.$$

From  $\|f'_2(x)\| \gg \frac{1}{q_1}$ , we have  $|\sum e(qf_2(n))| \leq q_1$ . Thus,

$$S \ll q \frac{\log Q}{\log X} q_1 + \frac{X}{\log^2 X} \ll q \frac{\log Q}{\log X} \exp((\log \log X)^3) + \frac{X}{\log^2 X} \ll \frac{X}{\log^2 X}.$$

This concludes the proof of Theorem 3.5.  $\square$

**Remark 3.1.** Note that Theorem 3.5 implies (and provides a new rather short proof of) the classical result of Rhin [Rh], which states that if  $f(x)$  is a polynomial with at least one coefficient other than the constant term irrational, then  $(f(p))_{p \in \mathcal{P}}$  is uniformly distributed mod 1. If one takes Rhin's theorem for granted, then the proof of Theorem 3.5 can be made shorter. Indeed, using summation by parts with  $a_n = e(qf_1(n))$  and  $b_n = \Lambda(n)e(qf_2(n))$  we get

$$\begin{aligned} \sum_{n=1}^N \Lambda(n) e(qf_1(n) + qf_2(n)) &= \\ \sum_{n=1}^{N-1} [e(qf_1(n)) - e(qf_1(n+1))] \sum_{m=1}^n \Lambda(m) e(qf_2(m)) &+ e(qf_1(N)) \sum_{m=1}^N \Lambda(m) e(qf_2(m)). \end{aligned}$$

Note that

$$|e(qf_1(n)) - e(qf_1(n+1))| \leq 2\pi q |f_1(n) - f_1(n+1)| \leq \frac{2\pi q C}{n}.$$

Denote  $c_n = (e(qf_1(n)) - e(qf_1(n+1))) \sum_{m=1}^n \Lambda(m) e(qf_2(m))$ . Then, by Rhin's theorem,

$$|c_n| \ll \left| \frac{1}{n} \sum_{m=1}^n \Lambda(m) e(qf_2(m)) \right| \rightarrow 0.$$

Therefore, for any non-zero integer  $q$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \Lambda(n) e(qf(n)) = 0,$$

so we are done.

We are now in the position to prove the main theorem in this section:

*Proof of Theorem 3.1.* We need to prove that (2) and (3) are equivalent. Let us first show that (3)  $\Rightarrow$  (2): we can write  $f(x) = f_1(x) + f_2(x)$ , where

- $f_1(x)$  is of the type  $x^{l+}$  for some non-negative integer  $l$  or  $\lim_{x \rightarrow \infty} f_1(x) = c$  for some  $c \in \mathbb{R}$ ,
- $f_2(x) \in \mathbb{R}[x]$ .



Case I.  $l \geq 1$ : the result follows from Theorems 3.3 and 3.4.

Case II.  $l = 0$ : If  $\lim_{x \rightarrow \infty} \left| \frac{f_1(x)}{\log x} \right| = \infty$ , the result follows from Theorems 3.2 and 3.4. Otherwise,  $\lim_{x \rightarrow \infty} |xf'_1(x)| = \lim_{x \rightarrow \infty} \left| \frac{f_1(x)}{\log x} \right| \leq C$  for some  $C$ . Then  $f_2(x) - f_2(0)$  has at least one irrational coefficient, so the result follows from Theorem 3.5.

Case III. If  $\lim_{x \rightarrow \infty} f_1(x) = c$  for some  $c \in \mathbb{R}$ , then, again, the result follows from Theorem 3.5.

Now let us show that (2)  $\Rightarrow$  (3): Suppose not. There exists a function  $f(x)$  such that  $(f(p))_{p \in \mathcal{P}}$  is u.d. mod 1 and  $\lim_{x \rightarrow \infty} \frac{f(x) - P(x)}{\log x} = A (\neq \pm\infty)$  for some  $P(x) \in \mathbb{Q}[x]$ . Find an integer  $q$  such that  $qP(x) \in \mathbb{Z}[x]$  and let  $g(x) = qf(x) - qP(x)$ . Then  $(g(p))_{p \in \mathcal{P}}$  is u.d. mod 1 since  $g(p) \equiv qf(p) \pmod{1}$ , thus for any non-zero integer  $h$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{\pi(N)} \sum_{p \leq N} e(hg(p)) = 0.$$

We have  $\lim_{x \rightarrow \infty} \frac{g(x)}{\log x} \neq \pm\infty$ , so  $|g'(x)| \ll \frac{1}{x}$ . The argument which was utilized in the course of the proof of Theorem 3.2 to establish the formula (3.5) gives us the following estimates for any non-zero integer  $h$

$$\begin{aligned} \sum_{p \leq N} e(hg(p)) &= \sum_{N_0 < n \leq N} \Lambda_1(n) e(hg(n)) + O\left(\frac{\pi(N)}{\log N}\right) \quad \left(N_0 = \frac{N}{\log N}\right) \\ &= \sum_{N_0 < n \leq N} \frac{1}{\log n} e(hg(n)) + O\left(\frac{\pi(N)}{\log N}\right) \\ &= \sum_{n \leq N} \frac{1}{\log n} e(hg(n)) + O\left(\frac{\pi(N)}{\log N}\right). \end{aligned}$$

Then

$$\lim_{N \rightarrow \infty} \frac{\log N}{N} \sum_{n \leq N} \frac{e(hg(n))}{\log n} = 0.$$

Using summation by parts, we get

$$\left| \sum_{n \leq N} e(hg(n)) \right| = \left| \sum_{n \leq N} \frac{e(hg(n))}{\log n} \log n \right| \ll \log N \max_{N_0 \leq N} \left| \sum_{n \leq N_0} \frac{e(hg(n))}{\log n} \right|.$$

Thus

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} e(hg(n)) = 0,$$

which implies that  $(g(n))_{n \in \mathbb{N}}$  is u.d. mod 1. By the result of Boshernitzan [Bos] alluded to in the introduction, this implies  $\lim_{x \rightarrow \infty} \frac{f(x) - P(x)}{\log x} = \pm\infty$ , which contradicts the above assumption that  $A \neq \pm\infty$ . We are done.  $\square$

#### 4. APPLICATIONS

For a given Hardy field  $H$ , let  $\mathbf{H}$  be the set of all subpolynomial functions  $\xi(x) \in H$  such that

$$\text{either } \lim_{x \rightarrow \infty} \frac{\xi(x)}{x^{l+1}} = \lim_{x \rightarrow \infty} \frac{x^l}{\xi(x)} = 0 \text{ for some } l \in \mathbb{N}, \text{ or } \lim_{x \rightarrow \infty} \frac{\xi(x)}{x} = \lim_{x \rightarrow \infty} \frac{\log x}{\xi(x)} = 0.$$

#### 4.1. Some corollaries of Theorem 3.1.

In this subsection we collect some consequences of results in previous sections, which will be utilized in subsequent subsections for derivation of ergodic and combinatorial applications.

**Proposition 4.1.** *Let  $P_1(x), \dots, P_m(x)$  be polynomials such that  $(P_1(n), \dots, P_m(n))_{n \in \mathbb{N}}$  is u.d. mod 1 in  $\mathbb{T}^m$  and let  $\xi_1(x), \dots, \xi_k(x) \in \mathbf{H}$  such that  $\sum_{i=1}^k b_i \xi_i(x) \in \mathbf{H}$  for any  $(b_1, \dots, b_k) \in \mathbb{Z}^k \setminus \{(0, 0, \dots, 0)\}$ . Then  $(P_1(p), \dots, P_m(p), \xi_1(p), \dots, \xi_k(p))_{p \in \mathcal{P}}$  is u.d. mod 1 in  $\mathbb{T}^{m+k}$ .*

*Proof.* By Weyl's criterion, it is enough to show that

$$(a_1 P_1(p) + \dots + a_m P_m(p) + b_1 \xi_1(p) + \dots + b_k \xi_k(p))_{p \in \mathcal{P}} \text{ is u.d. mod 1,}$$

where  $a_i, b_j \in \mathbb{Z}$  and at least one of  $a_i$  or  $b_j$  is not 0, which follows from Theorem 3.4.  $\square$

**Remark 4.1.** Here are some examples of  $\xi_1(x), \dots, \xi_k(x) \in \mathbf{H}$  satisfying the condition in Proposition 4.1.

- If  $\xi_1(x), \dots, \xi_k(x) \in \mathbf{H}$  have *different growth rates*, that is,

$$\lim_{x \rightarrow \infty} \frac{\xi_i(x)}{\xi_j(x)} = \pm \infty \quad \text{or} \quad 0 \quad \forall i \neq j,$$

then  $\sum_{i=1}^k b_i \xi_i(x) \in \mathbf{H}$  for any  $(b_1, \dots, b_k) \in \mathbb{R}^k \setminus \{(0, 0, \dots, 0)\}$ .

- $\xi_1(x) = \sqrt{x} + \log^2 x$  and  $\xi_2(x) = 2\sqrt{x} + \log^2 x$  do not have different growth rates, but they satisfy the condition in Proposition 4.1.

**Proposition 4.2.** *Let  $\xi_1(x), \dots, \xi_k(x) \in \mathbf{H}$  such that  $\sum_{i=1}^k b_i \xi_i(x) \in \mathbf{H}$  for any  $(b_1, \dots, b_k) \in \mathbb{R}^k \setminus \{(0, \dots, 0)\}$ . Let  $g(x) = \sum_{i=1}^k \alpha_i [\xi_i(x)]$ , where  $\alpha_1, \dots, \alpha_k$  be real numbers and let  $P(x) = a_0 + a_1 x + \dots + a_t x^t$  be a polynomial. Then,*

- (i) *If  $(\alpha_1, \dots, \alpha_k) \neq (0, \dots, 0)$ , then*

$$\lim_{N \rightarrow \infty} \frac{1}{\pi(N)} \sum_{p \leq N} e(g(p)) = 0.$$

- (ii) *If one of  $a_i (1 \leq i \leq t)$  or  $\alpha_i (1 \leq i \leq k)$  is irrational, then  $(g(p) + P(p))_{p \in \mathcal{P}}$  is u.d. mod 1.*

*Proof.* Without loss of generality we can assume that all  $\alpha_i$  are non-zero. Let  $p_n$  be the  $n$ -th prime number. Theorem 3.1 implies that

- (a)  $(\sum_{i=1}^k b_i \xi_i(p_n))_{n \in \mathbb{N}}$  is u.d. mod 1 for any  $(b_1, b_2, \dots, b_k) \neq (0, 0, \dots, 0)$ .
- (b)  $(\sum_{i=1}^k b_i \xi_i(p_n) + \sum_{j=0}^t c_j p_n^j)_{n \in \mathbb{N}}$  is u.d. mod 1 if  $(b_1, b_2, \dots, b_k) \neq (0, 0, \dots, 0)$  or one of the  $c_j$  is irrational.

Reordering  $\alpha_i$ , if needed, we will assume that there exists a non-negative integer  $m$  such that  $\alpha_i \notin \mathbb{Q}$  for  $i \leq m$  and  $\alpha_i \in \mathbb{Q}$  for  $i \geq m+1$ . For rational  $\alpha_i (i \geq m+1)$ , one has  $q \in \mathbb{N}$  such that  $\alpha_i = \frac{d_i}{q}$  with  $d_i \in \mathbb{Z}$ .

Let us first prove (i). Define  $f_i(x, y) = e(x - \{y\}\alpha_i)$  if  $\alpha_i$  is irrational and define  $f_i(x) = e(\alpha_i x)$  if  $\alpha_i$  is rational. Then  $e(g(p_n))$  is the product of  $f_i(\alpha_i \xi_i(p_n), \xi_i(p_n)) (i \leq m)$  and  $f_i([\xi_i(p_n)]) (i \geq m+1)$ . Note that (a) implies that for any  $u \in \mathbb{N}$ ,

$$\left( \alpha_1 \xi_1(p_n), \xi_1(p_n), \alpha_2 \xi_2(p_n), \xi_2(p_n), \dots, \alpha_m \xi_m(p_n), \xi_m(p_n), \frac{\xi_{m+1}(p_n)}{u}, \frac{\xi_{m+2}(p_n)}{u}, \dots, \frac{\xi_k(p_n)}{u} \right)$$

is u.d. mod 1 in  $\mathbb{T}^{m+k}$ . Thus,

$$(\alpha_1 \xi_1(p_n), \xi_1(p_n), \alpha_2 \xi_2(p_n), \xi_2(p_n), \dots, \alpha_m \xi_m(p_n), \xi_m(p_n), [\xi_{m+1}(p_n)], [\xi_{m+2}(p_n)], \dots, [\xi_k(p_n)])$$

is u.d. mod 1 in  $\mathbb{T}^{2m} \times \mathbb{Z}_q^{k-m}$ . So (i) follows.

Now let us prove (ii). If all  $\alpha_i = 0$ , then (ii) follows from Theorem 3.5. (It also follows from Rhin's theorem that  $P(p_n)$  is u.d. mod 1 ([Rh]).) Otherwise, without loss of generality we can assume that all  $\alpha_i \neq 0$ . We want to show that for any non-zero  $r \in \mathbb{Z}$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e(rP(p_n) + rg(p_n)) = 0. \quad (4.1)$$

Suppose that one of the  $a_i (i \geq 1)$  is irrational. Let  $f_0(x) = e(rx)$ ,  $f_i(x, y) = e(r(x - \alpha_i\{y\})) (1 \leq i \leq m)$ , and  $f_i(x) = e(r\frac{d_i}{q}x) (i \geq m+1)$ . Then  $e(rP(p_n) + rg(p_n))$  is the product of  $f_0(P(p_n))$ ,  $f_i(\alpha_i \xi_i(p_n), \xi_i(p_n)) (1 \leq i \leq m)$  and  $f_i([\xi_i(p_n)]) (i \geq m+1)$ . From (b), for any  $u \in \mathbb{N}$ ,

$$\left( P(p_n), \alpha_1 \xi_1(p_n), \xi_1(p_n), \alpha_2 \xi_2(p_n), \xi_2(p_n), \dots, \alpha_m \xi_m(p_n), \xi_m(p_n), \frac{\xi_{m+1}(p_n)}{u}, \frac{\xi_{m+2}(p_n)}{u}, \dots, \frac{\xi_k(p_n)}{u} \right)$$

is u.d. mod 1 in  $\mathbb{T}^{m+k+1}$ . Thus,

$$(P(p_n), \alpha_1 \xi_1(p_n), \xi_1(p_n), \alpha_2 \xi_2(p_n), \xi_2(p_n), \dots, \alpha_m \xi_m(p_n), \xi_m(p_n), [\xi_{m+1}(p_n)], [\xi_{m+2}(p_n)], \dots, [\xi_k(p_n)])$$

is u.d. mod 1 in  $\mathbb{T}^{2m+1} \times \mathbb{Z}_q^{k-m}$ , so (4.1) follows. Finally if all  $a_i$  are rational numbers, then  $\alpha_1$  is irrational. We can show (4.1) similarly by using that for any  $u \in \mathbb{N}$

$$\left( P(p_n) + \alpha_1 \xi_1(p_n), \xi_1(p_n), \alpha_2 \xi_2(p_n), \xi_2(p_n), \dots, \alpha_l \xi_m(p_n), \xi_m(p_n), \frac{\xi_{m+1}(p_n)}{u}, \frac{\xi_{m+2}(p_n)}{u}, \dots, \frac{\xi_k(p_n)}{u} \right)$$

is u.d. mod 1 in  $\mathbb{T}^{m+k}$ , so (4.1) follows.  $\square$

#### 4.2. Ergodic sequences.

In this subsection we deal with the sequences of the form  $\mathbf{d}_n = ([\xi_1(p_n)], \dots, [\xi_k(p_n)])$ , where  $\xi_1(x), \dots, \xi_k(x) \in \mathbf{H}$  and  $\sum_{i=1}^k b_i \xi_i(x) \in \mathbf{H}$  for any  $(b_1, \dots, b_k) \in \mathbb{R}^k \setminus \{(0, 0, \dots, 0)\}$ . The main result is that  $(\mathbf{d}_n)_{n \in \mathbb{N}}$  is an **ergodic sequence**: for any ergodic measure preserving  $\mathbb{Z}^k$ -action  $T = (T^{\mathbf{m}})_{(\mathbf{m} \in \mathbb{Z}^k)}$  on a probability space  $(X, \mathcal{B}, \mu)$  and for any  $f \in L^2$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f \circ T^{\mathbf{d}_n} = \int f d\mu \text{ in } L^2.$$

Recall the following version of the classical Bochner-Herglotz theorem.

**Theorem 4.1.** *Let  $U_1, \dots, U_k$  be commuting unitary operators on a Hilbert space  $\mathcal{H}$  and let  $f \in \mathcal{H}$ . Then there is a measure  $\nu_f$  on  $\mathbb{T}^k$  such that*

$$\langle U_1^{n_1} U_2^{n_2} \dots U_k^{n_k} f, f \rangle = \int_{\mathbb{T}^k} e^{2\pi i(n_1 \gamma_1 + \dots + n_k \gamma_k)} d\nu_f(\gamma_1, \dots, \gamma_k),$$

for any  $(n_1, n_2, \dots, n_k) \in \mathbb{Z}^k$ .

**Theorem 4.2.** *Let  $U_1, \dots, U_k$  be commuting unitary operators on a Hilbert space  $\mathcal{H}$ . Let  $\xi_1(x), \dots, \xi_k(x) \in \mathbf{H}$  such that  $\sum_{i=1}^k b_i \xi_i(x) \in \mathbf{H}$  for any  $(b_1, \dots, b_k) \in \mathbb{R}^k \setminus \{(0, 0, \dots, 0)\}$ . Then,*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N U_1^{[\xi_1(p_n)]} \dots U_k^{[\xi_k(p_n)]} f = f^*, \quad (4.2)$$

where  $f^*$  is the projection of  $f$  on  $\mathcal{H}_{inv} := \{f \in \mathcal{H} : U_i f = f \text{ for all } i\}$ .

*Proof.* We will use a Hilbert space splitting  $\mathcal{H} = \mathcal{H}_{inv} \oplus \mathcal{H}_{erg}$ , where

$$\begin{aligned} \mathcal{H}_{inv} &= \{f \in \mathcal{H} : U_i f = f \text{ for all } i = 1, 2, \dots, k\}, \\ \mathcal{H}_{erg} &= \{f \in \mathcal{H} : \lim_{N_1, \dots, N_k \rightarrow \infty} \left\| \frac{1}{N_1 \dots N_k} \sum_{n_1=0}^{N_1-1} \dots \sum_{n_k=0}^{N_k-1} U_1^{n_1} \dots U_k^{n_k} f \right\|_{\mathcal{H}} = 0\}. \end{aligned}$$

For  $f \in \mathcal{H}_{inv}$ ,  $U_1^{[\xi_1(p_n)]} \dots U_k^{[\xi_k(p_n)]} f = f$ . So we need to prove that for  $f \in \mathcal{H}_{erg}$ , the left hand side in (4.2) converges to 0. This follows from the Bochner-Herglotz theorem and Proposition 4.2(i):

$$\begin{aligned} & \left\| \frac{1}{N} \sum_{n=1}^N U_1^{[\xi_1(p_n)]} \dots U_k^{[\xi_k(p_n)]} f \right\|_{\mathcal{H}}^2 \\ &= \frac{1}{N^2} \sum_{m,n=1}^N \langle U_1^{[\xi_1(p_m)] - [\xi_1(p_n)]} \dots U_k^{[\xi_k(p_m)] - [\xi_k(p_n)]} f, f \rangle \\ &= \frac{1}{N^2} \sum_{m,n=1}^N \int e(i([\xi_1(p_m)] - [\xi_1(p_n)], \dots, [\xi_k(p_m)] - [\xi_k(p_n)]) \cdot \gamma) d\nu_f(\gamma) \\ &= \int \left| \frac{1}{N} \sum_{n=1}^N e(i([\xi_1(p_n)], \dots, [\xi_k(p_n)]) \cdot \gamma) \right|^2 d\nu_f(\gamma) \rightarrow 0, \end{aligned}$$

since  $f \in \mathcal{H}_{erg}$ , so  $\nu_f(\{0, 0, \dots, 0\}) = 0$ .  $\square$

**Corollary 4.1.** *Let  $\mathbf{d}_n = ([\xi_1(p_n)], \dots, [\xi_k(p_n)])$ , where  $\xi_1(x), \dots, \xi_k(x) \in \mathbf{H}$  and  $\sum_{i=1}^k b_i \xi_i(x) \in \mathbf{H}$  for any  $(b_1, \dots, b_k) \in \mathbb{R}^k \setminus \{(0, 0, \dots, 0)\}$ . Let  $T = (T^{\mathbf{m}})_{(\mathbf{m} \in \mathbb{Z}^k)}$  be an ergodic measure preserving  $\mathbb{Z}^k$ -action on a probability space  $(X, \mathcal{B}, \mu)$ . Then for any  $f \in L^2$*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f \circ T^{\mathbf{d}_n} = \int f d\mu \text{ in } L^2.$$

**Theorem 4.3.** *Let  $\xi_1(x), \dots, \xi_k(x) \in \mathbf{H}$  such that  $\sum_{i=1}^k b_i \xi_i(x) \in \mathbf{H}$  for any  $(b_1, \dots, b_k) \in \mathbb{R}^k \setminus \{(0, 0, \dots, 0)\}$ . Let  $L : \mathbb{Z}^k \rightarrow \mathbb{Z}^m$  be a linear transformation and  $(\psi_1(n), \dots, \psi_m(n)) = L([\xi_1(n)], \dots, [\xi_k(n)])$ . Let  $T_1, T_2, \dots, T_m$  be commuting, invertible measure preserving transformations on a probability space  $(X, \mathcal{B}, \mu)$ . Then, for any  $A \in \mathcal{B}$  with  $\mu(A) > 0$ , one has*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mu(A \cap T_1^{-\psi_1(p_n)} \dots T_m^{-\psi_m(p_n)} A) \geq \mu^2(A).$$

*Proof.* By Theorem 4.1 there exists a positive measure  $\nu$  on  $\mathbb{T}^m$  such that

$$\mu(A \cap T_1^{-n_1} \dots T_m^{-n_m} A) = \int e^{2\pi i(n_1 \gamma_1 + \dots + n_m \gamma_m)} d\nu(\gamma_1, \dots, \gamma_m).$$

Also one can see that  $\nu\{(0, \dots, 0)\} \geq \mu(A)^2$ . Let  $L^t : \mathbb{T}^m \rightarrow \mathbb{T}^k$  be defined by  $\mathbf{k} \cdot L^t(\mathbf{x}) = L(\mathbf{k}) \cdot \mathbf{x}$ , where  $\mathbf{k} \in \mathbb{Z}^k$  and  $\mathbf{x} \in \mathbb{T}^m$ . Denote by  $\nu'$  be the image of  $\nu$  under  $L^t$ . Then

$$\hat{\nu}'(\mathbf{k}) = \hat{\nu}(L(\mathbf{k})) \quad \text{and} \quad \nu'\{(0, \dots, 0)\} \geq \nu\{(0, \dots, 0)\}.$$

Moreover, by Proposition 4.2,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \hat{\nu}'([\xi_1(p_n)], \dots, [\xi_k(p_n)]) = \nu'\{(0, \dots, 0)\}.$$

Thus,

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mu(A \cap T_1^{-\psi_1(p_n)} \dots T_m^{-\psi_m(p_n)} A) &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^n \hat{\nu}(\psi_1(p_n), \dots, \psi_m(p_n)) \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \hat{\nu}'([\xi_1(p_n)], \dots, [\xi_k(p_n)]) = \nu'\{(0, \dots, 0)\} \geq \nu\{(0, \dots, 0)\} \geq \mu^2(A). \quad \square \end{aligned}$$

The next result follows from Theorem 4.3 with the help of Furstenberg's correspondence principle. (See Proposition 1.1.)

**Corollary 4.2.** *Let  $\psi_1(x), \dots, \psi_k(x)$  be as in Theorem 4.3. If  $E \subset \mathbb{Z}^k$  with  $d^*(E) > 0$ , then there exists a prime  $p$  such that  $(\psi_1(p), \dots, \psi_k(p)) \in E - E$ . Moreover,*

$$\liminf_{N \rightarrow \infty} \frac{|\{p \leq N : (\psi_1(p), \dots, \psi_k(p)) \in E - E\}|}{\pi(N)} \geq d^*(E)^2.$$

#### 4.3. Nice $FC^+$ sets.

Before stating the main result of this subsection, we will define some relevant notions. For  $\mathbf{d} = (d_1, d_2, \dots, d_k) \in \mathbb{Z}^k$ , we write  $|\mathbf{d}| = \max_{1 \leq i \leq k} |d_i|$ . Let  $D$  be an infinite subset of  $\mathbb{Z}^k$ . We write  $D = \{\mathbf{d}_n : n \in \mathbb{N}\}$  with the convention that  $\mathbf{d}_{n_1} \neq \mathbf{d}_{n_2}$  for  $n_1 \neq n_2$  and  $|\mathbf{d}_n|$  is non-decreasing.

##### Definition 4.1.

- (1) A set  $D \subset \mathbb{Z}^k$  is a **set of recurrence** if given any measure preserving  $\mathbb{Z}^k$ -action  $T = (T^{\mathbf{m}})_{(\mathbf{m} \in \mathbb{Z}^k)}$  on a probability space  $(X, \mathcal{B}, \mu)$  and any set  $A \in \mathcal{B}$  with  $\mu(A) > 0$ , there exists  $\mathbf{d} \in D$  ( $\mathbf{d} \neq 0$ ) such that

$$\mu(A \cap T^{-\mathbf{d}} A) > 0.$$

- (2) A set  $D \subset \mathbb{Z}^k$  is a **set of nice recurrence** if given any measure preserving  $\mathbb{Z}^k$ -action  $T = (T^{\mathbf{m}})_{(\mathbf{m} \in \mathbb{Z}^k)}$  on a probability space  $(X, \mathcal{B}, \mu)$ , any set  $A \in \mathcal{B}$  with  $\mu(A) > 0$  and any  $\epsilon > 0$ , we have

$$\mu(A \cap T^{-\mathbf{d}} A) \geq \mu^2(A) - \epsilon$$

for infinitely many  $\mathbf{d} \in D$ .

**Definition 4.2** (cf. Definition 1.2.1 in [BL]). A subset  $D$  of  $\mathbb{Z}^k \setminus \{0\}$  is a **van der Corput set** (**vdC set**) if for any family  $(u_{\mathbf{n}})_{\mathbf{n} \in \mathbb{Z}^k}$  of complex numbers of modulus 1 such that

$$\forall \mathbf{d} \in D, \lim_{N_1, \dots, N_k \rightarrow \infty} \frac{1}{N_1 \cdots N_k} \sum_{\mathbf{n} \in \prod_{i=1}^k [0, N_i)} u_{\mathbf{n} + \mathbf{d}} \overline{u_{\mathbf{n}}} = 0,$$

we have

$$\lim_{N_1, \dots, N_k \rightarrow \infty} \frac{1}{N_1 \cdots N_k} \sum_{\mathbf{n} \in \prod_{i=1}^k [0, N_i)} u_{\mathbf{n}} = 0.$$

**Definition 4.3.** An infinite subset  $D$  of  $\mathbb{Z}^k$  is a **nice  $FC^+$  set** if for any positive finite measure  $\sigma$  on  $\mathbb{T}^k$ ,

$$\sigma(\{(0, 0, \dots, 0)\}) \leq \limsup_{|\mathbf{d}| \rightarrow \infty, \mathbf{d} \in D} |\hat{\sigma}(\mathbf{d})|.$$

It is known that every nice  $FC^+$  set is a vdC set and a set of nice recurrence (see Section 3.5 in [BL] or Remark 4 in [BKMST]).

Let  $P_1(x), \dots, P_l(x) \in \mathbb{Z}[x]$  with  $P_i(0) = 0$  for all  $1 \leq i \leq l$  and let  $\xi_1(x), \dots, \xi_k(x) \in \mathbf{H}$  such that  $\sum_{i=1}^k b_i \xi_i(x) \in \mathbf{H}$  for any  $(b_1, \dots, b_k) \in \mathbb{Z}^k \setminus \{(0, 0, \dots, 0)\}$ . Let  $L$  be a non-zero linear map  $\mathbb{Z}^{l+k} \rightarrow \mathbb{Z}^m$ . Denote

$$\begin{aligned} D_{-1} &= \{(p-1, (p-1)^2, \dots, (p-1)^l, [\xi_1(p)], \dots, [\xi_k(p)]) \mid p \in \mathcal{P}\}, \\ D_1 &= \{(p+1, (p+1)^2, \dots, (p+1)^l, [\xi_1(p)], \dots, [\xi_k(p)]) \mid p \in \mathcal{P}\}. \end{aligned} \quad (4.3)$$

The main result in this subsection is following.

**Theorem 4.4.**  $D_1$  and  $D_{-1}$  are nice  $FC^+$  sets in  $\mathbb{Z}^m$ , and so they are vdC sets and also sets of nice recurrence.

Let  $L : \mathbb{Z}^{l+k} \rightarrow \mathbb{Z}^m$  be a non-zero linear map and let  $\mathbf{D}_i = L(D_i)$  ( $i = \pm 1$ ). Note that  $\mathbf{D}_i$  is an infinite set since  $(x \pm i), (x \pm i)^2, \dots, (x \pm i)^l, \xi_1(x), \dots, \xi_k(x)$  are linearly independent over  $\mathbb{Z}$ .

**Corollary 4.3.** Let  $P_1(x), \dots, P_l(x) \in \mathbb{Z}[x]$  with  $P_i(0) = 0$  for all  $1 \leq i \leq l$ . Then the following sets are nice  $FC^+$  sets:

$$\begin{aligned} &\{(P_1(p-1), P_2(p-1), \dots, P_l(p-1), [\xi_1(p)], \dots, [\xi_k(p)]) \mid p \in \mathcal{P}\}, \\ &\{(P_1(p+1), P_2(p+1), \dots, P_l(p+1), [\xi_1(p)], \dots, [\xi_k(p)]) \mid p \in \mathcal{P}\}. \end{aligned}$$

Note that for any  $(c_1, \dots, c_l, d_1, \dots, d_k) \in \mathbb{Z}^k \setminus \{(0, 0, \dots, 0)\}$ ,

$$|c_1(x-i) + \dots + c_l(x-i)^l + d_1[\xi_1(x)] + \dots + d_l[\xi_k(x)]|$$

is eventually increasing to  $\infty$ , so there exists a finite set  $F \subset D_i$ ,  $(0, \dots, 0) \notin L(D_i \setminus F)$ . Therefore the following result implies that in order to prove Theorem 4.4, it is enough to show that  $D_1$  and  $D_{-1}$  are nice  $FC^+$  sets.

**Lemma 4.1** (cf. [BL] Corollary 1.15). Let  $D \subset \mathbb{Z}^d$  be a nice  $FC^+$  set. Let  $L : \mathbb{Z}^d \rightarrow \mathbb{Z}^e$  be a linear transformation such that  $(0, 0, \dots, 0) \notin L(D)$ . Then  $L(D)$  is a nice  $FC^+$  set.

*Proof.* The proof is analogous to the proof of Theorem 4.3. Let  $\nu$  be a positive measure on  $\mathbb{T}^e$ . Define  $L^t : \mathbb{T}^e \rightarrow \mathbb{T}^d$  by  $\mathbf{k} \cdot L^t(\mathbf{x}) = L(\mathbf{k}) \cdot \mathbf{x}$ , where  $\mathbf{k} \in \mathbb{Z}^d$  and  $\mathbf{x} \in \mathbb{T}^e$ . Let  $\nu'$  be the image of  $\nu$  under  $L^t$ . Then

$$\hat{\nu}'(\mathbf{d}) = \hat{\nu}(L(\mathbf{d})) \quad \text{and} \quad \nu'\{(0, \dots, 0)\} \geq \nu\{(0, \dots, 0)\}.$$

Then

$$\nu\{(0, \dots, 0)\} \leq \nu'\{(0, \dots, 0)\} \leq \limsup_{\mathbf{d} \in D, |\mathbf{d}| \rightarrow \infty} |\hat{\nu}'(\mathbf{d})| = \limsup_{\mathbf{k} \in L(D), |\mathbf{k}| \rightarrow \infty} |\hat{\nu}(\mathbf{k})|. \quad \square$$

We will utilize the following lemmas to show that  $D_1$  and  $D_{-1}$  are nice  $FC^+$  sets.

**Lemma 4.2.** *Let  $\xi_1(x), \dots, \xi_k(x) \in \mathbf{H}$  such that  $\sum_{i=1}^k b_i \xi_i(x) \in \mathbf{H}$  for any  $(b_1, \dots, b_k) \in \mathbb{Z}^k \setminus \{(0, \dots, 0)\}$ . Let  $q \in \mathbb{N}$  and let  $t$  be an integer with  $0 \leq t \leq q-1$  and  $(t, q) = 1$ . Then  $(\xi_1(p), \dots, \xi_k(p))$  is u.d. mod 1 along  $p \in (t + q\mathbb{Z}) \cap \mathcal{P}$ .*

*Proof.* For  $(a_1, a_2, \dots, a_k) \neq (0, 0, \dots, 0)$  in  $\mathbb{Z}^k$ , let  $f(x) = \sum_{i=1}^k a_i \xi_i(x)$ . Define  $A_N = \{p \leq N : p \equiv t \pmod{q}\}$ . Then

$$\begin{aligned} \frac{1}{|A_N|} \sum_{\substack{p \leq N \\ p \equiv t \pmod{q}}} e(f(p)) &= \frac{1}{|A_N|} \sum_{p \leq N} e(f(p)) \frac{1}{q} \sum_{j=1}^q e\left(\frac{(p-t)j}{q}\right) \\ &= \frac{\pi(N)}{|A_N|} \frac{1}{q} \sum_{j=1}^q \frac{1}{\pi(N)} \sum_{p \leq N} e\left(f(p) + \frac{j}{q}(p-t)\right) \end{aligned}$$

and the result follows from the fact that  $\lim_{N \rightarrow \infty} \frac{|A_N|}{\pi(N)} = \frac{1}{\phi(q)}$  and  $(f(p) + \frac{j}{q}(p-t))_{p \in \mathcal{P}}$  is u.d. mod 1.  $\square$

**Lemma 4.3** (Lemma 4.1 in [BKMST]). *Let  $D \subset \mathbb{Z}^k$ . For each  $q \in \mathbb{N}$ , define*

$$D^{(q!)} := \{\mathbf{d} = (d_1, d_2, \dots, d_k) \in D : q! \text{ divides } d_i \text{ for all } 1 \leq i \leq k\}.$$

*Suppose that, for every  $q$ , there exists a sequence  $(\mathbf{d}^{q,n})_{n \in \mathbb{N}}$  in  $D^{(q!)}$  such that*

- (i)  $(|\mathbf{d}^{q,n}|)_{n \in \mathbb{N}}$  is non-decreasing and
- (ii) for any  $\mathbf{x} = (x_1, \dots, x_k) \in \mathbb{R}^k$ , if one of  $x_i$  is irrational, the sequence  $(\mathbf{x} \cdot \mathbf{d}^{q,n})_{n \in \mathbb{N}}$  is uniformly distributed mod 1.

*Then  $D$  is a nice  $FC^+$  set.*

*Proof of Theorem 4.4.* We will prove that  $D_1$  is a nice  $FC^+$  set. The proof for  $D_{-1}$  is analogous. Let

$$\mathbf{d}_n = (p_n + 1, \dots, (p_n + 1)^l, [\xi_1(p_n)], \dots, [\xi_k(p_n)]).$$

Let  $D_1^{(q!)} := \{\mathbf{d} = (d_1, d_2, \dots, d_{l+k}) \in D_1 : q! \text{ divides } d_i \text{ for } 1 \leq i \leq l+k\}$ .

Let us first show that  $D_1^{(q!)}$  has positive relative density in  $D_1$ . Consider the partition  $\mathcal{P} = \bigcup_{(t, q!) = 1} ((t + q!\mathbb{Z}) \cap \mathcal{P})$ . The relative density of  $(t + q!\mathbb{Z}) \cap \mathcal{P}$  in  $\mathcal{P}$  is  $\frac{1}{\phi(q!)}$ . Now, if  $p \in (t + q!\mathbb{Z}) \cap \mathcal{P}$ , the pair of conditions

$$q!|(t+1)^{c_i} \ (1 \leq i \leq l) \text{ and } 0 \leq \left\{ \frac{\xi_i(p)}{q!} \right\} < \frac{1}{q!} \ (1 \leq i \leq k)$$

is equivalent to  $(P_1(p+1), \dots, P_l(p+1), [\xi_1(p)], \dots, [\xi_k(p)]) \in D_1^{(q!)}$ . Then  $D_1^{(q!)}$  has positive relative density in  $D_1$  since  $\left(\frac{\xi_1(p)}{q!}, \dots, \frac{\xi_k(p)}{q!}\right)$  is uniformly distributed mod 1 in  $\mathbb{T}^k$  along the increasing sequence of primes  $p \in t + q!\mathbb{Z}$ .

Now let  $\mathbf{x} = (x_1, x_2, \dots, x_{l+k})$ , where one of  $x_i$  is irrational. We need to prove that for any non-zero integer  $m$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{|\{n \leq N : \mathbf{d}_n \in D_1^{(q^l)}\}|} \sum_{n \leq N, \mathbf{d}_n \in D_1^{(q^l)}} e(m(\mathbf{d}_n \cdot \mathbf{x})) = 0. \quad (4.4)$$

Then, using Lemma 2.8,

$$\begin{aligned} & \frac{1}{|\{n \leq N : \mathbf{d}_n \in D_1^{(q^l)}\}|} \sum_{n \leq N, \mathbf{d}_n \in D_1^{(q^l)}} e(m(\mathbf{d}_n \cdot \mathbf{x})) \\ &= \frac{1}{|\{n \leq N : \mathbf{d}_n \in D_1^{(q^l)}\}|} \sum_{n \leq N} e(m(\mathbf{d}_n \cdot \mathbf{x})) \frac{1}{(q^l)^{l+k}} \sum_{j_1=1}^{q^l} \cdots \sum_{j_{l+k}=1}^{q^l} e\left(\mathbf{d}_n \cdot \left(\frac{j_1}{q^l}, \dots, \frac{j_{l+k}}{q^l}\right)\right) \\ &= \frac{N}{|\{n \leq N : \mathbf{d}_n \in D_1^{(q^l)}\}|} \frac{1}{(q^l)^{l+k}} \sum_{j_1=1}^{q^l} \cdots \sum_{j_{l+k}=1}^{q^l} \frac{1}{N} \sum_{n \leq N} e\left(\mathbf{d}_n \cdot (m\mathbf{x} + \left(\frac{j_1}{q^l}, \dots, \frac{j_{l+k}}{q^l}\right))\right). \end{aligned}$$

Then (4.4) holds from Proposition 4.2.  $\square$

**Corollary 4.4.** *If  $E \subset \mathbb{Z}^m$  with  $d^*(E) > 0$ , then for any  $\epsilon > 0$*

$$R(E, \epsilon) := \{\mathbf{d} \in \mathbf{D}_i : d^*(E \cap E - \mathbf{d}) \geq d^*(E)^2 - \epsilon\}$$

*is infinite.*

#### 4.4. Uniform distribution and sets of recurrence.

Let  $\mathbf{D}_i = L(D_i)$  ( $i = \pm 1$ ) as in the previous section, that is,

$$D_{-1} = \{(p-1, \dots, (p-1)^l, [\xi_1(p)], \dots, [\xi_k(p)]) \mid p \in \mathcal{P}\},$$

$$D_1 = \{(p+1, \dots, (p+1)^l, [\xi_1(p)], \dots, [\xi_k(p)]) \mid p \in \mathcal{P}\},$$

where

- $L : \mathbb{Z}^{l+k} \rightarrow \mathbb{Z}^m$  is a non-zero linear map.
- $\xi_1(x), \dots, \xi_k(x) \in \mathbf{H}$  such that  $\sum_{i=1}^k b_i \xi_i(x) \in \mathbf{H}$  for any  $(b_1, \dots, b_k) \in \mathbb{Z}^k \setminus \{(0, 0, \dots, 0)\}$ .

**Theorem 4.5.** *Enumerate the elements of  $\mathbf{D}_i$ , ( $i = \pm 1$ ), as follows:*

$$\mathbf{d}_{n,i} = L((p_n + i), \dots, (p_n + i)^l, [\xi_1(p_n)], \dots, [\xi_k(p_n)]), \quad n = 1, 2, \dots$$

For each  $r \in \mathbb{N}$ , let  $\mathbf{D}_i^{(r)} = \mathbf{D}_i \cap \bigoplus_{j=1}^m r\mathbb{Z}$  and enumerate the elements of  $\mathbf{D}_i^{(r)}$  by

$(\mathbf{d}_{n,i}^{(r)})$  such that  $|\mathbf{d}_{n,i}^{(r)}|$  is non-decreasing.<sup>4</sup> Let  $(T^{\mathbf{d}})_{\mathbf{d} \in \mathbb{Z}^m}$  be a measure preserving  $\mathbb{Z}^m$ -action on a probability space  $(X, \mathcal{B}, \mu)$ . Then

(i)  $\mathbf{D}_i$  is an “averaging” set of recurrence:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mu(A \cap T^{-\mathbf{d}_{n,i}} A) > 0. \quad (4.5)$$

<sup>4</sup>Note that, for any  $r$  in  $\mathbb{N}$ , the sequence  $|\mathbf{d}_{n,i}^{(r)}|$  is eventually increasing.



(ii) For any  $A \in \mathcal{B}$  with  $\mu(A) > 0$  and for any  $\epsilon > 0$ , there exists  $r \in \mathbb{N}$  such that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mu(A \cap T^{-\mathbf{d}_{n,i}^{(r)}} A) \geq \mu(A)^2 - \epsilon. \quad (4.6)$$

Moreover,  $\{\mathbf{d} \in \mathbf{D}_i : \mu(A \cap T^{-\mathbf{d}} A) \geq \mu^2(A) - \epsilon\}$  has positive lower relative density in  $\mathbf{D}_i$ . Hence,  $\mathbf{D}_i$  is a set of nice recurrence.

*Proof.* Let us prove this for the case that  $L$  is an identity map on  $\mathbb{Z}^{l+k}$  and  $i = 1$  first. To simplify the notation we use  $\mathbf{d}_n$  and  $\mathbf{d}_n^{(r)}$  instead of  $\mathbf{d}_{n,i}$  and  $\mathbf{d}_{n,i}^{(r)}$ .

For any  $\mathbb{Z}^{l+k}$ -action  $T$ , there are commuting invertible measure preserving transformations  $T_1, \dots, T_{l+k}$  such that  $T^{\mathbf{m}} = T_1^{m_1} \cdots T_{l+k}^{m_{l+k}}$  for  $\mathbf{m} = (m_1, m_2, \dots, m_{l+k})$ .

First we will show that the limits in (4.5) and (4.6) exist. By Theorem 4.1, there exists a measure  $\nu$  on  $\mathbb{T}^{k+l}$  such that

$$\mu(A \cap T^{-\mathbf{n}} A) = \int 1_A(x) T^{\mathbf{n}} 1_A(x) d\mu(x) = \int_{\mathbb{T}^{l+k}} e(\mathbf{n} \cdot \gamma) d\nu(\gamma).$$

Thus, it is sufficient to show that for every  $\gamma$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e(\mathbf{d}_n \cdot \gamma) \quad \text{and} \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e(\mathbf{d}_n^{(r)} \cdot \gamma) \quad (4.7)$$

exist. By (2.14), denoting  $A_N = \{n \leq N : \mathbf{d}_n \in D_1^{(r)}\}$ ,

$$\begin{aligned} & \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e(\mathbf{d}_n^{(r)} \cdot \gamma) \\ &= \lim_{N \rightarrow \infty} \frac{1}{|A_N|} \sum_{n=1}^N e(\mathbf{d}_n \cdot \gamma) \left( \prod_{i=1}^l \frac{1}{r} \sum_{j_i=1}^r e\left(\frac{(p_n+1)^{j_i}}{r}\right) \right) \left( \prod_{i=1}^k \frac{1}{r} \sum_{j_{l+i}=1}^r e\left(\frac{[\xi_i(p_n)]j_{l+i}}{r}\right) \right) \\ &= \lim_{N \rightarrow \infty} \frac{N}{|A_N|} \frac{1}{r^{k+l}} \sum_{j_1=1}^r \cdots \sum_{j_{l+k}=1}^r \frac{1}{N} \sum_{n=1}^N e\left(\mathbf{d}_n \cdot \left(\gamma + \left(\frac{j_1}{r} + \cdots + \frac{j_{l+k}}{r}\right)\right)\right). \end{aligned}$$

Using the same argument as in the proof of Theorem 4.4, we can show that relative density of  $D_1^{(r)}$  in  $D_1$  is positive, so we only need to show that  $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e(\mathbf{d}_n \cdot \gamma)$  exists for every  $\gamma$ .

From Proposition 4.2, if  $\gamma \notin \mathbb{Q}^{l+k}$ , then  $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e(\mathbf{d}_n \cdot \gamma) = 0$ . If  $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_{l+k}) \in \mathbb{Q}^{l+k}$ , then we can find a common denominator  $q \in \mathbb{N}$  for  $\gamma_1, \dots, \gamma_{l+k}$  such that  $\gamma_i = \frac{a_i}{q}$  for each  $i$ . We claim that the following limit exists:

$$\lim_{N \rightarrow \infty} \frac{1}{\pi(N)} \sum_{\substack{p \equiv t \pmod{q} \\ p \leq N}} e\left(\sum_{j=1}^k [\xi_j(p)] \frac{a_{l+j}}{q}\right).$$

This follows from two observations:

- (a)  $\{p \in \mathcal{P} : p \equiv t \pmod{q}\}$  has a density  $\frac{1}{\phi(q)}$  in  $\mathcal{P}$  for  $(t, q) = 1$ .
- (b)  $([\xi_1(p)], \dots, [\xi_k(p)])$  is u.d. in  $\mathbb{Z}_q^k$  along  $p \in t + q\mathbb{Z}$  for  $(t, q) = 1$  since  $\left(\frac{\xi_1(p)}{q}, \dots, \frac{\xi_k(p)}{q}\right)$  is u.d. mod 1 in  $\mathbb{T}^k$  along  $p \in t + q\mathbb{Z}$  from Lemma 4.2.

Thus, we have

$$\begin{aligned}
\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e(\mathbf{d}_n \cdot \gamma) &= \lim_{N \rightarrow \infty} \frac{1}{\pi(N)} \sum_{p \leq N} e \left( \sum_{i=1}^l P_i(p+1) \frac{a_i}{q} + \sum_{j=1}^k [\xi_j(p)] \frac{a_{l+j}}{q} \right) \\
&= \lim_{N \rightarrow \infty} \frac{1}{\pi(N)} \sum_{\substack{(t,q)=1 \\ 0 \leq t \leq q-1}} \sum_{\substack{p \equiv t \pmod{q} \\ p \leq N}} e \left( \sum_{i=1}^l P_i(p+1) \frac{a_i}{q} + \sum_{j=1}^k [\xi_j(p)] \frac{a_{l+j}}{q} \right) \\
&= \sum_{\substack{(t,q)=1 \\ 0 \leq t \leq q-1}} e \left( \sum_{i=1}^l P_i(t+1) \frac{a_i}{q} \right) \lim_{N \rightarrow \infty} \frac{1}{\pi(N)} \sum_{\substack{p \equiv t \pmod{q} \\ p \leq N}} e \left( \sum_{j=1}^k [\xi_j(p)] \frac{a_{l+j}}{q} \right).
\end{aligned}$$

Now let us show (ii). Consider the following Hilbert space splitting for  $L^2(X) = \mathcal{H} = \mathcal{H}_{rat} \oplus \mathcal{H}_{tot}$  for  $T_1, \dots, T_{l+k}$ , where

$$\mathcal{H}_{rat} = \overline{\{f \in \mathcal{H} : \text{there exists non-zero } (m_1, m_2, \dots, m_{l+k}) \in \mathbb{Z}^{l+k}, T_i^{m_i} f = f \text{ for all } i\}},$$

$$\mathcal{H}_{tot} = \{f \in \mathcal{H} : \text{for any non-zero } (m_1, m_2, \dots, m_{l+k}) \in \mathbb{Z}^{l+k}$$

$$\lim_{N_1, \dots, N_k \rightarrow \infty} \left\| \frac{1}{N_1 \cdots N_{l+k}} \sum_{n_1=0}^{N_1-1} \cdots \sum_{n_k=0}^{N_k-1} T_1^{m_1 n_1} \cdots T_{l+k}^{m_{l+k} n_{l+k}} f \right\|_{\mathcal{H}} = 0\}.$$

Let  $1_A = f + g$ , where  $f \in \mathcal{H}_{rat}$  and  $g \in \mathcal{H}_{tot}$ . Note that  $\mathcal{H}_{rat} = \overline{\bigcup_{q=1}^{\infty} \mathcal{H}_q}$ , where  $\mathcal{H}_q = \{f : T_i^{q^l} f = f \text{ for } i = 1, 2, \dots, l+k\}$ .

For any  $\epsilon > 0$ , there exists  $\mathbf{a} = (a_1, \dots, a_{l+k}) \in \mathbb{Z}^{l+k}$  and  $f_{\mathbf{a}} \in \mathcal{H}_{rat}$  such that  $T^{\mathbf{a}} f_{\mathbf{a}} = f_{\mathbf{a}}$ ,  $\|f_{\mathbf{a}} - f\| < \epsilon/2$  and  $\int f_{\mathbf{a}} d\mu = \mu(A)$ . Choose  $r$  such that  $a_i | r$  for all  $i$ . Consider

$$\frac{1}{N} \sum_{n=1}^N \mu(A \cap T^{-\mathbf{d}_n^{(r)}} A) = \frac{1}{N} \sum_{n=1}^N \int f(x) T^{\mathbf{d}_n^{(r)}} f(x) d\mu(x) + \frac{1}{N} \sum_{n=1}^N \int g(x) T^{\mathbf{d}_n^{(r)}} g(x) d\mu(x).$$

For  $f \in \mathcal{H}_{rat}$ ,

$$\begin{aligned}
\int f(x) T^{\mathbf{d}_n^{(r)}} f(x) d\mu(x) &= \langle f_{\mathbf{a}}, f_{\mathbf{a}} \rangle + \langle f_{\mathbf{a}}, T^{\mathbf{d}_n^{(r)}} (f - f_{\mathbf{a}}) \rangle + \langle f - f_{\mathbf{a}}, T^{\mathbf{d}_n^{(r)}} f \rangle \\
&\geq \mu^2(A) - \epsilon,
\end{aligned}$$

Also note that  $(\mathbf{d}_n^{(r)} \cdot \gamma)$  is u.d. mod 1 for  $\gamma \notin (\mathbb{Q}/\mathbb{Z})^{l+k}$ . Hence,

$$\frac{1}{N} \sum_{n=1}^N \int g(x) T^{\mathbf{d}_n^{(r)}} g(x) d\mu(x) = \int \frac{1}{N} \sum_{n=1}^N e(\mathbf{d}_n^{(r)} \cdot \gamma) d\nu(\gamma) \rightarrow 0,$$

since  $\nu(\mathbb{Q}/\mathbb{Z})^{l+k} = 0$  because  $g \in \mathcal{H}_{tot}$ . Then,

$$\frac{1}{N} \sum_{n=1}^N \mu(A \cap T^{-\mathbf{d}_n^{(r)}} A) \geq \mu(A)^2 - \epsilon,$$

so  $\{\mathbf{d} \in D_1 : \mu(A \cap T^{-\mathbf{d}} A) \geq \mu^2(A) - \epsilon\}$  has positive lower relative density in  $D_1$ .

Now it remains to show (i). Choose  $\epsilon$  so small that  $\mu^2(A) - \epsilon \geq \mu^2(A)/2$ . Since  $(\mathbf{d}_n^{(r)})$  has positive relative density, say  $\delta$ , we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mu(A \cap T^{-\mathbf{d}_n} A) \geq \delta \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mu(A \cap T^{-\mathbf{d}_n^{(r)}} A) \geq \frac{\delta}{2} \mu^2(A).$$

The proof for  $i = -1$  is completely analogous. By the argument in the proof of Theorem 4.3 we can prove the theorem for any  $\mathbf{D}_i = L(D_i)$ .  $\square$

**Corollary 4.5.** *If  $E \subset \mathbb{Z}^m$  with  $d^*(E) > 0$ , then for any  $\epsilon > 0$ ,*

$$\{\mathbf{d} \in \mathbf{D}_i : d^*(E \cap E - \mathbf{d}) \geq d^*(E)^2 - \epsilon\}$$

*has positive lower relative density in  $\mathbf{D}_i$  for  $i = \pm 1$ . Furthermore,*

$$\liminf_{N \rightarrow \infty} \frac{|\{p \leq N : L(p-1, \dots, (p-1)^l, [\xi_1(p)], \dots, [\xi_k(p)]) \in E - E\}|}{\pi(N)} > 0.$$

$$\liminf_{N \rightarrow \infty} \frac{|\{p \leq N : L(p+1, \dots, (p+1)^l, [\xi_1(p)], \dots, [\xi_k(p)]) \in E - E\}|}{\pi(N)} > 0.$$

**Remark 4.2.** Let  $P_1(x), \dots, P_l(x) \in \mathbb{Z}[x]$  with  $P_i(0) = 0$  for all  $1 \leq i \leq l$  and let  $\xi_1(x), \dots, \xi_k(x) \in \mathbf{H}$  such that  $\sum_{i=1}^k b_i \xi_i(x) \in \mathbf{H}$  for any  $(b_1, \dots, b_k) \in \mathbb{Z}^k \setminus \{(0, 0, \dots, 0)\}$ . By choosing a linear map  $L$  appropriately, the sets  $D_{-1}$  and  $D_1$  in Theorem 4.5 and Corollary 4.5 become

$$\{(P_1(p-1), P_2(p-1), \dots, P_l(p-1), [\xi_1(p)], \dots, [\xi_k(p)]) \mid p \in \mathcal{P}\},$$

$$\{(P_1(p+1), P_2(p+1), \dots, P_l(p+1), [\xi_1(p)], \dots, [\xi_k(p)]) \mid p \in \mathcal{P}\}$$

respectively.

ACKNOWLEDGMENTS. The authors would like to thank Donald Robertson for careful reading of the manuscript and many useful comments.

## REFERENCES

- [BH] V. Bergelson and I. J. Håland Knutson, Weak mixing implies weak mixing of higher orders along tempered functions. *Ergodic Theory and Dynamical Systems* **29** (2009), no. 5, 1375-1416.
- [BKMST] V. Bergelson, G. Kolesnik, M. Madritsch, Y. Son and R. Tichy, Uniform distribution of prime powers and sets of recurrence and van der Corput sets in  $\mathbb{Z}^k$ . *Israel Journal of Mathematics* **201** (2014), no.2, 729-760.
- [BL] V. Bergelson and E. Lesigne, Van der Corput sets in  $\mathbb{Z}^d$ , *Colloquium Mathematicum* **110** (2008), no. 1, 1-49.
- [BMc] V. Bergelson and R. McCutcheon, An ergodic IP polynomial Szemerédi theorem, *Mem. Amer. Math. Soc.* **146** (2000), no. 695.
- [Bos] M. Boshernitzan, Uniform distribution and Hardy fields. *J. Anal. Math.* **62** (1994), 225-240.
- [DT] M. Drmota and R. F. Tichy, *Sequences, discrepancies and applications*. Lecture Notes in Mathematics, 1651. Springer-Verlag, Berlin, 1997.
- [F] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. d'Analyse Math.* **31** (1977), 204-256.
- [GK] S.W. Graham and G. Kolesnik, *Van der Corput's method of exponential sums*, London Mathematical Society Lecture Note Series, **126**. Cambridge University Press, Cambridge, 1991.
- [Har1] G. H. Hardy, Properties of logarithmico-exponential functions, *Proc. London Math. Soc.* **10** (1912), 54-90.

- [Har2] G. H. Hardy, *Orders of infinity*, (Cambridge Tracts in Math. and Math. Phys. vol. 12), 2nd edition, Cambridge, 1924.
- [KM] T. Kamae and M. Mendès France, Van der Corput's difference theorem, *Israel J. Math* **31** (1977), 335-342.
- [Kar] A. A. Karatsuba, *Basic analytic number theory*, Springer-Verlag, 1992.
- [KN] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Pure Appl. Math., Wiley-Interscience, 1974.
- [Mo] H. L. Montgomery, *Ten Lectures on the Interface Between Analytic Number Theory and Harmonic Analysis*, CBMS Reg. Conf. Ser. Math. **84**, Amer. Math. Soc., 1994.
- [Rh] G. Rhin, Sur la répartition modulo 1 des suites  $f(p)$ , *Acta. Arith.* **23** (1973), 217-248.
- [Sa1] A. Sárközy, On difference sets of sequences of integers. I. *Acta Math. Acad. Sci. Hungar.* **31** (1978), no. 1-2, 125-149.
- [Sa2] A. Sárközy, On difference sets of sequences of integers. II. *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **21** (1978), 45-53.
- [Sa3] A. Sárközy, On difference sets of sequences of integers. III. *Acta Math. Acad. Sci. Hungar.* **31** (1978), no. 3-4, 355-386.
- [vdC] J. G. van der Corput, Diophantische Ungleichungen I. Zur Gleichverteilung modulo Eins, *Acta Math.* **56** (1931), 373-456.
- [Va] R. C. Vaughan, An elementary method in prime number theory. *Acta Arith.* **37** (1980), 111-115.
- [Vin] I. M. Vinogradov, *The method of trigonometric sums in the theory of numbers* (Russian), Izdat. "Nauka", Moscow, 1971.

(V. Bergelson) DEPARTMENT OF MATHEMATICS, OHIO STATE UNIVERSITY, COLUMBUS, OH 43210, USA

*E-mail address:* `vitaly@math.ohio-state.edu`

(G. Kolesnik) DEPARTMENT OF MATHEMATICS, CALIFORNIA STATE UNIVERSITY, LOS ANGELES, CA 90032, USA

*E-mail address:* `gkolesnik@sbcglobal.net`

(Y. Son) FACULTY OF MATHEMATICS AND COMPUTER SCIENCE, WEIZMANN INSTITUTE OF SCIENCE, REHOVOT, 7610001, ISRAEL

*E-mail address:* `younghwan.son@weizmann.ac.il`